PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

United States Patent and Trademark Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 10 June 1999 (10.06.99)	
International application No. PCT/GB98/03449	**Applicant's or agent's file reference** M L/C320.01/G
International filing date (day/month/year) 16 November 1998 (16.11.98)	**Priority date** (day/month/year) 14 November 1997 (14.11.97)
Applicant MURPHY, Jonathan, Paul et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

17 May 1999 (17.05.99)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer Lazar Joseph Panakal
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 33/74, 33/58

A1

(11) International Publication Number: **WO 99/26069**

(43) International Publication Date: 27 May 1999 (27.05.99)

(21) International Application Number: PCT/GB98/03449

(22) International Filing Date: 16 November 1998 (16.11.98)

(30) Priority Data:
9723955.2 14 November 1997 (14.11.97) GB

(71) Applicant *(for all designated States except US)*: GENERIC BIOLOGICALS LIMITED [GB/GB]; 8 Centre One, Old Sarum Park, Lysander Way, Salisbury, Wiltshire SP4 6BU (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MURPHY, Jonathan, Paul [GB/GB]; 37 Bishops Orchard, East Hagbourne, Oxfordshire OX11 9JS (GB). ATKINSON, Anthony [GB/GB]; Twingley, Mill Corner, Winterbourne Gunner, Salisbury, Wiltshire SP4 6JJ (GB).

(74) Agent: KEITH W. NASH & CO.; 90–92 Regent Street, Cambridge CB2 1DP (GB).

(81) Designated States: JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMPROVEMENTS IN OR RELATING TO DETECTION OF MOLECULES IN SAMPLES

```
F P T I P L S R L F D N A M L R A H R L H Q L A F 25
D T Y Q E W E E A Y I P K E Q K Y S F L Q N P Q T 50
S L C F S E S I P T P S N R E E T Q Q K S N L E L 75
L R I S L L L I Q S W L E P V Q F L R S V W A N S 100
L V Y G A S D S N V Y D L L K D L E E G I Q T L M 125
G R L E D G S P R T G Q I P K Q T Y S K F D T N S 150
H N D D A L L K N Y G L L Y C F R K D M D K V E T 175
F L R I V Q C R S V E G S C G F                   191
```

(57) Abstract

Disclosed is a method of detecting the presence in a sample of a polypeptide exogenously administered to a mammalian subject from whom the sample is obtained, and distinguishing between such an exogenously administered polypeptide and a naturally–occurring endogenous polypeptide present in the sample; the method comprising obtaining a sample from the subject; and subjecting the sample to analysis of fluorescence at a suitable wavelength; wherein the exogenously administered polypeptide is tagged with a greater or lesser amount of fluorescence activity, relative to the untagged endogenous polypeptide, at the wavelength(s) analysed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Title: Improvements in or Relating to Detection of Molecules in Samples

Field of the Invention

This invention relates to tagged molecules (distinguishable from untagged, but otherwise identical, molecules), methods of preparing tagged molecules, nucleic acid sequences and constructs encoding tagged molcules, and a method of distinguishing between tagged and untagged (but otherwise identical) molecules.

In particular, the present invention relates to a method of tagging a protein with a therapeutically acceptable tag which enables detection of the tagged protein administered exogenously to humans, bovines or other animals where the same (but untagged) protein is produced endogenously, and a method of detecting and differentiating the tagged protein over the endogenous protein. In particular, the method is suitable for application to human growth hormone (hGH), to enable differential detection of exogenously administered hGH in humans, for example, to determine whether hGH is being administered unlawfully for its performance enhancing effects.

Background of the Invention

Previously, the usual method of differentiating exogenously administered protein from the endogenous one has been to tag the exogenous protein with radioactive labels. Because of the hazards of radioactivity, radioactively tagged proteins are administered to patients over short periods of time in controlled conditions and under medical supervision. Further, radioactive labels are not therapeutically acceptable since they are intrusive to the biological system in which such tagged proteins are administered. Other tagging methods tend to alter the biological function of the protein molecule and therefore, are no longer suitable for therapeutic use. Such prior art tagging methods are therefore limited to controlled research uses and do not have widespread cost effective commercial applications.

Some amino acids, for example tryptophan (W) and tyrosine (Y) in particular, are natural fluorophores, which fluoresce when appropriately stimulated. This fluorescence can be detected and measured with standard prior art fluorescence detection techniques. Proteins

which contain such fluorophores in their amino acid sequence may also fluoresce when appropriately stimulated. The level of fluorescence can be crudely related to the number of fluorophores in the protein. The fluorescent yield of any fluorophore is sensitive to its local environment such that, for example, there may be a difference between its fluorescence in an aqueous and a hydrophobic environment. Waldman *et al* (1987 Biochem. Biophys. Acta *931*, 66-71; 1988 Biochem. Biophys Res. Comm. *150* (2), 752-759), Corinne (1991 Biochemistry *30*, 1028-1036) and others have exploited this property to perform *in vitro* laboratory studies on conformational and structural changes of lactate dehydrogenase when, for example, substrate binding occurs. Waldman and Corrine have mutated lactate dehydrogenase to incorporate tryptophan residues at the substrate binding site. This technique is restricted to use as a research tool for conformational and structural studies of proteins *in vitro*, since often the full biological activity or structural conformation of the native protein is lost. Thus, such modified proteins are no longer suitable for therapeutic purposes and there is no disclosure or suggestion of pharmaceutical compositions comprising the mutated protein. Moreover, there is no disclosure or suggestion in the prior art that such mutations could form the basis for a method of distinguishing the altered compound from the naturally occurring compound.

WO 94/10200 discloses and is concerned with amino acid substitutions in somatotropin (i.e. Growth Hormone) which provide increased conformational and chemical stability.

There is no suggestion in WO 94/10200 that modifications can be made to Growth Hormone for the purpose of distinguishing between endogenous Growth Hormone present in a subject and exogenous Growth Hormone administered to the subject. A number of amino acid substitutions in somatotropin are disclosed or suggested in WO 94/10200 which, because of the natural fluorophore activity of the amino acid residues tryptophan and tyrosine (discussed above), result in a somatotropin molecule having an altered fluorescence activity relative to the wild type, unsubstituted molecule. Such substitutions include the following:

G40→Y (i.e. glycine substituted by tyrosine at residue number 40); F52→Y; W86→F, Y, L, I or V; F103→Y; I137→Y;

A reliable method for differentiating and detecting exogenously administered hGH is particularly desirable when attempting to monitor the pharmacokinetics and/or

pharmacodynamics of hGH, or to detect its unlawful administration by athletes and others who illicitly use hGH for improving their performance. Presently, standard detection methods (e.g. HPLC, ELISA), are used for measuring the total amount of hGH in an athletes' blood or urine samples, and by subtracting the expected levels referenced to the general population, estimations of elevated hGH levels can be made. However, as levels vary considerably between individuals, and exogenous levels fall rapidly with time, this is a very crude measurement. In addition, as the performance enhancing effects last much longer than the detectable transient elevated levels of hGH in these samples, unless samples are taken shortly after administration the technique does not give indisputable proof that exogenous hGH has or has not been used.

The present invention seeks to alleviate the above mentioned problems by tagging or modifying a protein (such as hGH) with a therapeutically acceptable tag which can be detected simply and can be differentiated from the endogenous protein present in a sample of cells, blood, urine or other body fluid. The invention has little or no effect on the biological activity of the protein, such that the modified protein can be administered therapeutically in the same manner as the unmodified protein. Thus, the modified or tagged protein can be safely prescribed by physicians for existing or new therapeutic purposes, and also economically manufactured commercially at substantially the same cost as the untagged protein.

A further advantage of the present invention is that although levels of the exogenous protein may drop rapidly after administration, the specificity for the tagged protein and high sensitivity of the detection method allow detection long after the exogenous protein has been administered. Thus, an abuser cannot claim abnormally elevated production of the endogenous protein, and unlawful use of the tagged protein can be detected. Additionally, the present invention allows the pharmacokinetics and/or pharmacodynamics of the tagged exogenous protein to be detected and monitored.

Therefore, it is an object of the present invention to provide a method for tagging proteins which method enables detection of the exogenous tagged protein over any endogenous polypeptide which may be present in a sample (e.g. such as blood or urine) taken from, for example, a human subject (e.g. an athlete) or other mammalian subject (e.g. domesticated

farm livestock).

It is another object of the present invention to provide a modified polypeptide molecule, such as hGH, tagged in a manner which is therapeutically acceptable. Further, the tagging method of the present invention enables the biological activity *per se* of a protein to remain substantially unaltered such that the therapeutic efficacy is maintained and the protein can be administered in a manner identical to or similar with the unmodified protein.

A further specific object of the present invention is to provide a modified hGH molecule substituted with tryptophans at strategic positions in the native hGH sequence.

Summary of the Invention

In a first aspect, the invention provides a method of detecting the presence in a sample of a polypeptide exogenously administered to a mammalian subject from whom the sample is obtained, and distinguishing between such an exogenously administered polypeptide and a naturally-occurring endogenous polypeptide present in the sample; the method comprising obtaining a sample from the subject; and subjecting the sample to analysis of fluorescence at a suitable wavelength; wherein the exogenously administered is tagged with a greater or lesser amount of fluorescence activity, relative to the untagged endogenous polypeptide, at the wavelength(s) analysed.

In a second aspect, the invention provides a composition for administration to a mammalian subject, the composition comprising a polypeptide and a physiologically acceptable carrier substance, characterised in that the polypeptide is tagged with a greater or lesser amount of fluorescent activity relative to an untagged polypeptide endogenously present in the subject, the tagged molecule thereby being distinguishable from the untagged molecule by analysis of the fluorescence characteristics of the respective molecules, excluding those compositions in which the tagged molecule is Growth Hormone and wherein the fluorescent tagging consists solely of one or more of the following substitutions in the tagged Growth Hormone: G40 → Y; F52 → Y; W86 → F, Y, L, I or V; F103 → Y; and I137 → Y.

The tagged molecule is a polypeptide, which may typically be administered to a mammalian subject to exert a beneficial effect (e.g. for clinical or veterinary reasons, or for reasons of animal husbandry). The mammalian subject will generally be human, but may also be a

domesticated animal, especially a farm animal such as a bovine, porcine or ovine animal. The tagged molecule will generally therefore be a therapeutic polypeptide (i.e. comprises five or more amino acid residues and has a desirable effect on the subject, with little or no undesired side effect, when administered in an appropriate dose) and will possess the same biological activity as, and normally be substantially identical (except for the tagging) to, a naturally-occurring polypeptide present in the subject, although where the tagged molecule is a recombinant polypeptide it may have additional slight differences relative to the naturally occurring polypeptide (e.g. to increase activity, or to increase stability, e.g. as taught in WO 94/10200). (The "biological activity" of the molecule is that activity by which the molecule exerts its beneficial effect on the subject e.g. stimulation of growth in the case of GH; or stimulation of erythrocyte production in the case of EPO.)

The molecule may be, for example, a pharmaceutical. A particularly preferred molecule is a mammalian growth hormone, especially human growth hormone (hGH), bovine growth hormone (bGH), or porcine growth hormone (pGH); or calcitonin; or erythropoietin (EPO). Accordingly it is preferred that any fluorophores present in the tagged molecule: (a) do not have any significant effect on the biological activity of the molecule; and (b) are essentially non-toxic (that is, any fluorophores present will not cause the tagged molecule to exhibit any toxicity for the subject when the molecule is administered at normal therapeutic doses). Accordingly, tryptophan or tyrosine and closely-related compounds are preferred fluorophores for use in tagged molecules in accordance with the invention. These have the additional advantage of being readily incorporated into polypeptide molecules.

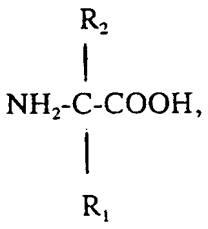
Advantageously, the tagged molecule is either deficient in, or comprises additional, fluorescent entities (fluorophores) relative to the untagged molecule. The tagging may therefore be "positive" (in which the tagged molecule comprises additional fluorophores) or "negative" (where the tagged molecule is deficient in fluorophores relative to the untagged molecule).

As explained above, the naturally occurring amino acid residues tryptophan (W) and, to a lesser extent, tyrosine (Y), possess natural fluorophore activity. Thus, if an "untagged" polypeptide comprises one or more tryptophan and/or tyrosine residues it may be fluorescent. Thus a tagged molecule, in accordance with the invention, may be distinguishable from an

untagged molecule by having additional fluorophores (especially if the untagged polypeptide comprises no, or very few, tryptophan or tyrosine residues and thus possesses no, or very little, intrinsic fluorescence). Alternatively, where the untagged molecule comprises a fluorophore (especially a plurality of fluorophores), the tagged molecule may be distinguishable by having fewer fluorophores than the untagged molecule.

Preferably, the tagged molecule comprises additional fluorophores present in amino acid residues or other compounds which are capable of forming a peptide bond, and thus are capable of being covalently incorporated into a polypeptide, either internally during synthesis of the polypeptide, and/or at the C-terminal after synthesis of the bulk of the polypeptide.

Conveniently the fluorophores additionally present in (or absent from) the tagged molecule (relative to the untagged molecule) are tyrosine and/or tryptophan residues, or a synthetic amino acid derivative wherein a fluorophore is covalently joined to an "amino acid" backbone, the synthetic derivative having the general formula



wherein R_1 comprises the fluorophore and R_2 is H, OH, halide or lower alkyl (C_1 to C_3, substituted or unsubstituted). The fluorophore R_1 may be a fluorophore which is present in a naturally occurring amino acid residue (e.g. the aromatic groups of tryptophan or tyrosine) or may be some other fluorophore (typically comprising a delocalised electron system, such as in an aromatic or heterocyclic ring). Such synthetic amino acid derivatives are already known in the art or can readily be prepared using standard organic chemistry techniques.

As a less preferable alternative to the tagged molecule comprising a different number of fluorophores (relative to the untagged molecule), the tagged molecule may comprise the fluorophores at different positions - the immediate chemical environment can affect the level of fluorescence of a fluorophore. Accordingly, the tagged molecule may not have a different number of fluorophores relative to the untagged molecule, but they may be of different

fluorescent activities and/or be differently disposed within the molecule so as to affect their fluorescence.

Where the tagged molecule is a polypeptide, tagging is conveniently accomplished by substituting a non-fluorescent amino acid present in the untagged molecule for an amino acid residue comprising a fluorophore (such as tryptophan, tryosine or a synthetic amino acid derivative), so as to increase the fluorescence of the tagged molecule relative to the untagged molecule.

With the benefit of the teaching of the present specification, and with the benefit of information otherwise readily available as common general knowledge, the person skilled in the art can, by routine trial and error, find appropriate amino acid residues which can be substituted, without substantially affecting the biological activity of the molecule. Conveniently, phenylalanine residues (F) or tyrosine residues (Y) can be replaced with tryptophan residues (W), which exhibit far greater fluorescence activity. Such substitutions are "conservative" and thus tend not to have any significant effect on the biological properties of a polypeptide. Further guidance for the person skilled in the art is given in the example below, which utilises principles which are generally applicable to any biologically active polypeptide.

The composition will normally comprise an effective amount of the tagged molecule, such that the biological activity thereof produces a demonstrable effect when administered to the subject. An "effective amount" is the amount of tagged molecule which results in the desired biological effect in the mammalian subject to which the composition is administered. The desired effect will, of course, depend on the identity of the tagged molecule: where the tagged molecule is EPO, for example, the desired effect is an increase in the number of erythrocytes per unit volume of blood in the subject. In some embodiments the composition will be essentially sterile, and suitable for delivery by means of injection (e.g. by trans-dermal, intravenous, intramuscular or subcutaneous routes). In other embodiments the composition will be in the form of a tablet, pill or capsule (e.g. enteric-coated capsules for slow release) for oral consumption.

Administration of the compositions of the invention into a mammalian subject may be performed according to known methods using any route effective to deliver the required

dosage to the subject. Modes of administration include those typically encountered for the species of choice. Because proteins in general are susceptible to degradation in the digestive system, injection is preferred via an intramuscular, transdermal or subcutaneous route. The use of sustained or prolonged release formulations or implants are also suitable modes. Generally, injection of a sustained release formulation is preferred.

The effective dosage range depends on the species, age, weight, and general health of the mammalian subject. These and other parameters which are needed to determine the effective dosage range for a given mammal are well within the purview of one skilled in the art. For instance, in bovines the effective amount of bovine GH (whether tagged or untagged) is in the range of 1.0 to 200 milligrams per animal per day. In pigs, for instance the effective amount of porcine GH is about 60 μg/kg/day.

The physiologically acceptable carrier may be a sterile liquid diluent where the composition is injected (e.g. saline, phosphate-buffered saline, or other aqueous buffer preparation). Where the composition is to be administered orally or transdermally, the carrier may be calcium carbonate, calcium sulphate or other substantially inert solid. Transdermal delivery by means of a needleless injection device may generally be preferred.

Methods of performing the fluorescence analysis may be entirely conventional and well known to those skilled in the art (e.g. spectrofluorimetry). The choice of method will depend in part on the manner in which the exogenous substance is tagged, and the characteristics of the fluorophore (if any) employed in the tagged molecule. For example, where the tagged molecule comprises fewer or more tryptophan residues than the untagged molecule, fluorescence analysis will typically be performed at about 297nm excitation.

Advantageously the sample is subjected to processing, prior to fluorescence analysis. to enrich or purify the endogenous and (if present) exogenous molecules in the sample. This improves the signal-to-noise ratio. Various methods of enrichment or purification may be employed, using one or more of the following techniques: centrifugation; HPLC; FPLC; affinity chromatography; immunoaffinity chromatography; heat treatment at 50-55°C for ten minutes (this is particularly appropriate for purification of growth hormone, which is relatively heat-stable - contaminating proteins will tend to be denatured, aggregate and precipitate, and so can be simply removed by centrifugation whilst the undenatured growth

hormone stays in solution); all of which are well known to those skilled in the art. The preferred method may depend, at least in part, on the identity of the endogenous and exogenous molecules.

The method defined immediately above is extremely useful in detecting the presence of exogenously administered molecules used illicitly by cyclists, athletes and others to improve performance. Very often, such molecules occur naturally (e.g. EPO, hGH, and the like) and are endogenous to the athlete's body, such that proving illicit use of performance-enhancing substances is very difficult. However, with the benefit of the present invention, such substances can be tagged, and thus made distinguishable over endogenous molecules synthesised naturally in the athlete's body.

Additionally the invention can be used to monitor the persistence of substances administered to the body. For example, the pharmacokinetics and/or pharmacodynamics of various drugs can readily be monitored by comparing fluorescence activities at different time points - this is particularly useful where the tagged drug is otherwise identical to an endogenous compound.

In a preferred embodiment, the tagged molecule is a polypeptide prepared using recombinant DNA technology. In such embodiments the method may additionally comprise the preparation of a nucleic acid sequence encoding the tagged molecule, the sequence being mutated relative to the wild type sequence encoding the untagged molecule. Typically the nucleic acid sequence encoding the tagged polypeptide comprises nucleotide substitutions (relative to the wild type sequence) so as to direct the expression of a polypeptide having one or more tryptophan residues not present in the untagged molecule or, less preferably, directing the expression of a polypeptide having fewer tryptophan residues than in the untagged molecule.

The nucleic acid sequence encoding the tagged molecule may be prepared, for example, by mutation of the wild-type sequence (e.g. by site-directed mutagenesis), by polymerase chain reaction (PCR), or by *de novo* synthesis (e.g. using an automated DNA synthesiser). All of these techniques are familiar and well-known to those skilled in the art and/or are readily obtained by reference to standard texts in the field (e.g. Sambrook *et al*, "Molecular Cloning, A Laboratory Manual" Cold Spring Harbor Laboratory Press, 1989).

Where the subject is a human, the sample may conveniently be a sample of body fluids, such as a blood, sweat, semen, urine, or saliva sample. Less preferably the sample may be a tissue sample comprising cells (e.g. skin scrapings from the buccal cavity, hair or the like). Where the subject is a domesticated farm animal, the sample may be taken from the animal before or after slaughter. Samples taken after slaughter conveniently include muscle tissue or other solid tissues taken from the carcass.

In another aspect of the present invention there is provided a tagged GH molecule comprising a tryptophan residue substituted for a phenylalanine residue present in a naturally-occurring molecule. In one embodiment, tryptophan is substituted at positions F31 and/or F97 in the amino acid sequence.

In a preferred embodiment, the tagged growth hormone comprises a tryptophan residue at one or more of positions 10, 31, 97, 160 or 176 (of which tryptophan residues at positions 31 and/or 97 are especially preferred). The tagged growth hormone molecule is preferably tagged hGH.

According to a still further aspect of the present invention there is provided a nucleic acid expression vector comprising substantially nucleotides 114-695 of the nucleic acid sequence shown in Figure 2. The CPG_2 signal sequence (nucleotides 39-113) is intended to direct the encoded polypeptide product to the bacterial periplasm - those skilled in the art will appreciate that the CPG_2 signal does not form an essential part of the vector, but is useful for expression in prokaryotes. Other signal sequences are well known to those skilled in the art and could be substituted for the CPG_2 signal sequence if desired. Thus the expression vector may be designed to cause expression in eukaryotes (e.g. mammalian tissue culture, fungal or yeast cultures) or in prokaryotes (bacterial cultures). In a particular embodiment the expression vector is a prokaryotic expression system, preferably comprising the vector pMTLhGHm described below.

The invention will now be described by way of illustrative examples and with reference to the accompanying drawings, in which:-

Figure 1 shows the primary amino acid sequence (Seq. ID No. 1) of native hGH protein;

Figure 2 shows a nucleic acid sequence (Seq. ID No. 2) encoding a tagged hGH molecule

for use in the method of present invention:

Figure 3 shows the primary amino acid sequence (Seq. ID No. 3) of the tagged hGH molecule encoded by the nucleic acid sequence of Figure 2;

Figure 4 is a schematic representation of the nucleic acid construct pMTLhGHm used to express a tagged polypeptide in accordance with the invention;

Figure 5 shows the amino acid sequence of human calcitonin (Seq. ID No. 4) - the sequence is shown in the orientation N terminal→C terminal, but the C terminal residue includes a naturally occurring amide group (as a post-translational modification);

Figure 6 shows the amino acid sequence of human growth hormone releasing factor (HGHRF) (Seq. ID No. 5) - the sequence is shown in the orientation N terminal→C terminal, but the C terminal residue includes a naturally occurring amide group (as a post-translational modification);

Figures 7A and 7B show the amino acid sequence of the A and B chains respectively of human insulin (Seq. ID Nos. 6 and 7);

Figure 8 shows the amino acid sequence of human Erythropoietin (Seq. ID No. 8); and

Figure 9 shows the amino acid sequence of human Interleukin 2 (Seq. ID No. 9).

Example 1 - Construction of an enhanced fluorescent form of hGH.

Amino acids can be generally classified into 4 main classes depending on their R groups: (1) non-polar or hydrophobic R groups; (2) neutral (uncharged) polar R groups; ((3) positively charged R groups; and (4) negatively charged R groups.

Although within any single class there is considerable variation in the size, shape, and properties of the R groups, certain amino acids show similar properties and can often be substituted without dramatically changing the protein conformation or biological activity. It has been suggested that there is a 70-80% chance of attaining a mutated protein with unchanged biological properties, by replacing any one phenylalanine (F) or tyrosine (Y) residue by tryptophan (W). However, it is not always possible accurately to predict the actual effect of such substitutions on the conformation and/or biological activity of a protein.

The extent of the effect or sensitivity of the protein to such substitution(s), will depend on the function of the target amino acid residue which is to be substituted. If the target amino acid is involved in catalysis or interacts with another residue, the protein will be sensitive to substitution. However, if the target amino acid is a scaffolding residue, the protein will be less sensitive to substitution by an amino acid with a similar R group.

In the native hGH amino acid sequence, there are twelve F, eight Y and one W residues (see Figure 1). In the method described below, techniques available in the art are used to establish which of the F and Y residues will have the least effect on the structure and biological activity of hGH, if substituted with W.

Determination:

Using the available 3D X-ray structures of hGH alone and hGH complexed with its receptor (both obtained from the Brookhaven data base) the positions of twenty potential substitutions were analysed in order to filter out the sensitive substitutions by Environmental Filtering which involves:-

1) Eliminating residues close to the surface of the protein:-

The inventors find that substitutions at surface sites should be avoided since the added hydrophobic character of the tryptophan residue sometimes gives rise to increased protein aggregation. Further, modifications to residues at the surface of a polypeptide are generally undesirable as they may (i) interfere with binding activity of the protein; and (ii) are more likely to create a new epitope which may be recognised as foreign by the immune system of the recipient. The following residues were shown to be surface residues and were therefore regarded as poor candidates for substitution in hGH:- F1; F25; Y35; Y42; F44; F54; F92; Y103; Y111; Y143; F146; F191.

2) Eliminating residues close to inter-protein surfaces: -

The remaining eight residues were determined to be buried in the protein conformational structure. Y164 and Y28 were determined to have close proximity to the hGH receptor glutamate residue and therefore likely to be critical in the interaction between hGH and its receptor. Thus, residues Y164 and Y28 are poor candidates for substitution.

3) Eliminating residues close to W86:-

It is known that fluorophores which are in close proximity to each other can interact through internal energy transfer (a blue shift in emission) thereby quenching the individual fluorescence of each fluorophore. F166 was determined to be in close proximity to W86, the single naturally occurring fluorophore in the native hGH protein. In order to avoid such potential quenching effects which could reduce the desired effect of fluorescence enhancement of the mutated protein, F166 was regarded as a poor candidate for substitution.

From the above analysis the inventors determined that the best candidates for W substitutions are: F31, F97, F10, F176 and/or Y160. Further analysis of the environmental position of these five residues within the native hGH revealed suitability for W substitution (see Table 1). In the example given below, F31 and F97 were selected for construction of the modified protein, and the other remaining residues are potentially suitable candidates also. It will be appreciated that other amino acids at suitable sites may be similarly substituted, and that, whilst substitution of F or Y residues is preferred, the invention is not so limited.

Table 1 shows the five tryptophan substitutions predicted as least likely to alter the hGH protein conformation and therefore least likely to affect biological activity.

RESIDUE	ENVIRONMENT	DISTANCE TO RECEPTOR SURFACE
F31	Hydrophobic cluster just below surface. Between helices.	Great
F97	Deep in a surface cleft	Great
F10	Hydrophobic surface slot	Adequately remote
F176	Buried, but close to W86	Adequately remote
Y160	Hydrophobic cluster just below the surface	About 0.6nm from receptor surface

The principles described above (i.e. avoiding substitution of residues which are surface-exposed or near functional sites such as active or allosteric sites of enzymes or receptor-

binding sites of ligands; and avoiding substitution of residues near other fluorescent residues) can be used to identify phenylalanine residues in any other biologically active molecule which are suitable for substitution by tryptophan, thereby allowing any desired polypeptide to be tagged, relative to a naturally-occurring endogenous polypeptide, so that the method of the invention can be applied very generally.

Example 2 - Construction of gene by substituting W for F31 and F97.

Optimised gene sequence: Using the empirically observed codon utilisation bias for highly expressed *E. coli* genes, the known DNA sequence coding for native hGH was re-designed incorporating, where possible, this *E. coli* codon bias whilst ensuring retention of the original translated protein sequence. The two substitutions (W31 and W97) were then incorporated into this optimised *E. coli* gene sequence. It will be appreciated by those skilled in the art that optimisation of codon bias for *E. coli* may not be desirable if the sequence is to be expressed in a host other than *E. coli*.

Figure 2 shows the nucleotide sequence used to encode the modified hGH. To facilitate expression and subsequent purification, the hGH coding sequence is preceded by a 75bp fragment of DNA derived from the carboxypeptidase G_2 (CPG_2) gene encoding the twenty-five amino acid signal peptide of CPG_2 (Minton *et al*, 1985 Gene *31*, 31-38). The CPG_2 signal peptide directs the expressed protein to the periplasm where the CPG_2 signal sequence is enzymatically cleaved, releasing the authentic hGH protein into the periplasm. Those skilled in the art will appreciate that the CPG_2 signal peptide sequence could be replaced with any one of a large number of functionally equivalent signal sequences from other sources, without substantially affecting the nature of the construct.

The amino acid sequence of the tagged hGH encoded by the nucleic acid sequence is shown in Figure 3; the tryptophan substitutions at positions 31 and 97 are shown in bold type.

Construction of the synthetic gene: This synthetic gene was constructed by standard chemical procedures (Wosnick *et al*, 1987 Gene *60* (1), 115-127) using double-stranded annealed pairs of 60-100 bp oligonucleotides with appropriately compatible sticky ends.

Cloning: Using standard techniques, the synthesised gene was restricted with *Nde* I and *Xho* I and cloned into the identical restriction sites of pMTL1015 (Chambers *et al*, 1986 Gene

68(1), 139-149) to produce pMTL hGHm (illustrated schematically in Figure 4). This vector directs expression of the synthetic gene under the control of the *mdh* promoter (Alldread *et al*, 1992 Gene 114(1), 139-143), and carries a selectable tetracycline resistance gene (Tc^R). Those skilled in the art will appreciate that the *mdh* promoter could be replaced with any one of a large number of functionally equivalent promoters from other sources, without substantially affecting the nature of the construct.

The new construct, pMTLhGHm, was transformed and subsequently expressed in an appropriate strain of *E. coli* (K-12 strain RV308, ATCC 31608) using standard procedures.

Production and Purification: The modified hGH protein (hGHf) may be produced in an industrial scale fermenter by methods well known to those skilled in the art. For example, a transformed *E. coli* culture containing pMTLhGHm may be grown up in aqueous media in a steel or other fermentation vessel conventionally aerated and agitated, in aqueous media at e.g. about 28-37°C and near neutral pH, supplied with appropriate nutrients such as glycerol, nitrogen sources such as ammonium sulphate, potassium sources such as potassium phosphate, trace elements, magnesium sulphate and the like. The plasmid pMTLhGHm carries tetracycline resistance as a selectable characteristic, so that selection pressure (i.e. inclusion in the medium of tetracycline at 12.5μg/ml) may be imposed to discourage competitive growth from wild-type organisms which lack the resistance characteristic (e.g. due to "segregation" of the plasmid during growth of the culture).

Upon completion of fermentation the cell suspension is centrifuged or the cellular solids otherwise collected from the broth and then lysed by physical or chemical means. Cellular debris is removed from supernatant and soluble hGHf isolated and purified.

HGHf may be purified from cell extracts using one or more of the following techniques: (i) polyethyleneimine fractionation; (ii) gel filtration chromatography on Sephacryl S-200; (iii) ion exchange chromatography on ToyoPearl Super Q 650m or CM Sephadex; (iv) hydrophobic chromatography using Phenyl-Sepharose; (v) ammonium sulphate and/or pH fractionation; (vi) selective heat enrichment; and (vii) affinity chromatography using antibody resins prepared from anti-hGH IgG isolated from immunosensitised animals or hybridomas; and desorbed under acid or slightly denaturing conditions. In particular, recombinant Growth Hormone may be purified from *E. coli* cultures according to the method disclosed in WO

87/00204 or EP 0 177 343.

Example 3 - Fluorescence detection.

In order to detect the hGHf by fluorescent measurements in samples from a mammalian subject to whom the hGHf has been administered, it is preferable to purify or enrich the sample (i.e. blood or urine) to reduce background fluorescent interference. This can be routinely accomplished by the use of a number of standard chromatographic techniques such as HPLC, FPLC, affinity chromatography, or immunoaffinity chromatography. Fluorescence may be increased by prior denaturation of the protein, for example by use of mild heat treatment and/or chaotropic agents (e.g. 1-6M Urea or guanidimium chloride).

W-fluorescence is measurable using standard techniques such as, for example, an SLM 8000 single photon counting spectrofluorometer. The purified sample is subjected to excitation around 297nm across a 2mm cell using a mercury-Xenon arc lamp and fluorescence detected around 345nm using a Mullard XP 2020Q rapid-response photomultiplier along a 1cm path at 90^0 to the excitation beam. Scattered light is excluded by cut-off filter (Schott 310) between the sample and photomultiplier.

An alternative embodiment of the invention can be envisaged, in which exogenous hGH is provided with reduced fluorescence relative to the naturally occurring molecule, for example by replacing W at position 86 with either F or Y.

It will be appreciated that the present invention has applications in other areas such as detection of exogenous proteins over the same protein produced endogenously, for example, measuring exogenous bovine growth hormone (bGH) which is administered to increase milk or meat production in cattle. Additionally the methods of the present invention can be used to detect abuse of such anabolic proteins in humans or in animals.

It will be further appreciated that the present invention is not limited to mammalian growth hormone proteins and can be equally successfully applied to other proteins including those which are also produced endogenously and those with therapeutic applications, such as calcitonin.

Example 4 - Human Calcitonin

Calcitonin (thyrocalcitonin) is an endogenous 32 amino acid peptide hormone produced by certain cells in the thyroid gland whose principle action is to lower the levels of calcium and phosphate in the blood. It is used clinically to treat several disorders such as hypercalcaemia and bone disorders such as Paget's disease and Osteoporosis. The amino acid sequence of calcitonin is illustrated in Figure 5, and included as Seq. ID No. 4 in the attached Sequence Listing.

Calcitonin may be negatively tagged (i.e. provided with reduced fluorescence) or positively tagged (i.e. provided with reduced fluorescence) relative to the naturally occurring molecule as follows:

To reduce fluorescence: replace Y 12 with L;

To increase fluorescence any one of the following substitutions may be performed: replace Y 12 with W; replace any F residue (located at positions 16, 19, and 22) with W; replace any two F residues (located at positions 16, 19, and 22) with W, preferably F 16 and F 22 (so as to avoid possible complications of "quenching" or other interference if fluorophores are too close together); replace any F residue (located at positions 16, 19, and 22) with Y, preferably F 22; replace any two F residues (located at positions 16, 19, and 22) with Y, preferably F 16 and F 22.

Example 5 - Human Growth Hormone Releasing Factor

Human Growth Hormone Releasing Factor (HGHRF) is an endogenous 44 amino acid peptide hormone that controls the release of human growth hormone. Consequently its clinical uses are similar to those for human growth hormone itself. The amino acid sequence of HGHRF is shown in Figure 6, and included as Seq. ID No. 5 in the attached sequence listing. HGHRF may be positively tagged (i.e. provided with increased fluorescence) relative to the naturally occurring molecule, by performing any one of the following substituions: replace any one of R 41, 42, or 43 with W; or replace both R 41 and R 43 with W.

Example 6 - Human Insulin

Human insulin is an endogenous hormone produced in the pancreas by the beta cells of the islets of Langerhans and is important for regulating the amount of glucose in the blood.

Lack of this hormone gives rise to diabetes mellitus, and as such insulin is used clinically to treat this condition. Mature insulin consists of two peptides, termed A and B, which are joined by two disulphide bridges: one between A chain C7 and B chain C7; and a second between A chain C20 and B chain C19. The sequence of the A and B chains of human insulin are shown in Figures 7A and 7B respectively, and are included as Seq. ID Nos. 6 and 7 in the attached sequence listing.

Human insulin may be negatively tagged or positively tagged, relative to the naturally occurring molecule, so as to be provided with reduced or increased fluorescence respectively, as described below:

To reduce fluorescence: replace any one or more Y residues (located at positions A 14; B 16 B 26) with either L or F;

To increase fluorescence: either; replace any F residue (located at positions B 24 and B 25) with W; replace any Y residue (located at positions A 14; B 16; and B 26) and either F residue (located at positions B 24 and B 25) with W.

Example 7 - Human Erythropoietin (EPO)

Human Erythropoietin is the principal endogenous factor responsible for the regulation of red blood cell production during steady-state conditions and for accelerating recovery of red blood cell mass following haemorrhage. As a result, EPO has important clinical uses where elevated levels of red blood cell expression is indicated. The amino acid sequence of EPO is shown in Figure 8, and is included as Seq. ID No. 8 in the attached sequence listing. EPO may conveniently be negatively tagged relative to naturally occurring EPO by replacing any one or more W residues (located at positions 51, 64 and 88) with F.

Example 8 - Human Interleukin 2 (IL-2)

Human Interleukin 2 is an endogenous factor produced and secreted primarily by activated T helper cells that acts as a paracrine factor driving the expansion of antigen specific cells and as a paracrine factor influencing the activity of a number of other cells including B cells, NK cells and LAK cells. Because of this central role of the IL-2/IL-2R system in mediation of the immune response, IL-2 has important diagnostic and therapeutic implications. For

example, IL-2 has shown promise as an anti-cancer drug by virtue of its ability to stimulate the proliferation and activities of tumour-attacking LAK and TIL cells. The amino acid sequence of human IL-2 is shown in figure 9 and is included as Seq. ID No. 9 in the attached sequence listing.

Human IL-2 may conveniently be negatively tagged or positively tagged (i.e. provided with reduced or increased fluorescence, respectively) relative to naturally occurring IL-2 as follows:

To reduce fluorescence: replace W 121 with either Y or F;

To increase fluorescence: either; replace any one or more F residues (located at positions 42, 44, 78, and 103) with W; or replace any one or more Y residues (located at positions 31, 45 and 107) with W.

CLAIMS

1. A method of detecting the presence in a sample of a polypeptide exogenously administered to a mammalian subject from whom the sample is obtained, and distinguishing between such an exogenously administered polypeptide and a naturally-occurring endogenous polypeptide present in the sample; the method comprising obtaining a sample from the subject; and subjecting the sample to analysis of fluorescence at a suitable wavelength; wherein the exogenously administered polypeptide is tagged with a greater or lesser amount of fluorescence activity, relative to the untagged endogenous polypeptide, at the wavelength(s) analysed.

2. A method according to claim 1, wherein the sample is subjected to processing, prior to fluorescence analysis, to enrich or purify the exogenous and endogenous molecules in the sample.

3. A method according to claim 1 or 2, wherein the sample is subjected to processing, prior to analysis, by one or more of the following: centrifugation; HPLC; FPLC; affinity chromatography; immunoaffinity chromatography; denaturation or heat treatment.

4. A method according to any one of claims 1, 2 or 3, wherein the sample is a sample of body fluid or tissue obtained from a human or other mammalian subject.

5. A method according to any one of the preceding claims, wherein the sample comprises one or more of the following: blood; saliva; sweat; urine; semen; tears.

6. A method according to any one of the preceding claims, wherein the tagged molecule has greater fluorescence activity, at the wavelength analysed, than the untagged molecule.

7. A method according to any one of the preceding claims, wherein the tagged molecule comprises one or more fluorophores not present in the untagged molecule.

8. A method according to claim 7, wherein a compound comprising a tagging fluorophore is incorporated in the tagged molecule by means of a peptide bond.

9. A method according to claim 7 or 8, wherein the fluorophore comprises tyrosine, tryptophan or a synthetic amino acid derivative.

10. A method according to any one of the preceding claims, wherein the tagged molecule comprises a tagged therapeutic polypeptide and/or tagged hormone.

11. A method according to any one of the preceding claims, wherein the tagged molecule comprises one of the following: a tagged human, bovine or porcine growth hormone; tagged calcitonin; tagged erythropoietin; tagged growth hormone releasing factor; tagged insulin; or tagged interleukin-2.

12. A method according to any one of the preceding claims, wherein the tagged molecule comprises growth hormone tagged with a tryptophan residue at one or more of positions 10, 31, 97, 160 or 176.

13. A composition for administration to a mammalian subject, the composition comprising a polypeptide and a physiologically acceptable carrier substance, characterised in that the polypeptide is tagged with a greater or lesser amount of fluorescent activity relative to an untagged polypeptide endogenously present in the subject, the tagged molecule thereby being distinguishable from the untagged molecule by analysis of the fluorescence characteristics of the respective molecules, excluding those compositions in which the tagged molecule is Growth Hormone and wherein the fluorescent tagging consists solely of one or more of the following substitutions in the tagged Growth Hormone: G40 → Y; F52 → Y; W86 → F, Y, L, I or V; F103 → Y; and I137 → Y.

14. A composition according to claim 13, wherein the tagged molecule comprises a number of tryptophan residues different from the number of tryptophan residues present in the untagged molecule, and the tagging is effected thereby.

15. A composition according to claim 13 or 14, wherein the tagged molecule comprises two or more tryptophan residues greater than the number of tryptophan residues present in the untagged molecule.

16. A composition according to any one of claims 13, 14 or 15, wherein the tagged molecule comprises a therapeutic polypeptide and/or hormone.

17. A composition according to any one of claims 13-16, wherein the tagged molecule comprises one of the following: tagged human, bovine or porcine growth hormone; tagged

calcitonin; tagged erythropoietin; tagged growth hormone releasing factor; tagged insulin; or tagged interleukin-2.

18. A composition according to any one of claims 13-17, wherein the tagged molecule comprises growth hormone tagged with a tryptophan residue at one or more of positions 10, 31, 97, 160 or 176.

19. A tagged growth hormone comprising a tryptophan residue substituted for a phenylalanine residue present in a naturally-occurring growth hormone molecule.

20. A tagged growth hormone comprising a tryptophan residue at one or more of positions 10, 31, 97, 160 or 176.

21. A tagged growth hormone comprising a tryptophan residue at position 31 and/or 97.

22. A nucleic acid sequence encoding a tagged growth hormone in accordance with any one of claims 19, or 21.

23. A nucleic acid expression construct comprising a nucleic acid sequence in accordance with claim 22.

24. A nucleic acid sequence comprising nucleotides 114-695 of the nucleic acid sequence shown in Figure 2.

25. A method substantially as hereinbefore defined.

26. A composition substantially as hereinbefore defined.

Fig.1.

F	P	T	I	P	L	S	R	L	**F**	D	N	A	M	L	R	A	H	R	L	H	Q	L	A	**F**	25
D	T	**Y**	Q	E	**F**	E	E	A	**Y**	I	P	K	E	Q	K	**Y**	S	**F**	L	Q	N	P	Q	T	50
S	L	C	**F**	S	E	S	I	P	T	P	S	N	R	E	E	T	Q	Q	K	S	N	L	E	L	75
L	R	I	S	L	L	L	I	Q	S	**W**	L	E	P	V	Q	**F**	L	R	S	V	**F**	A	N	S	100
L	V	**Y**	G	A	S	D	S	N	V	**Y**	D	L	L	K	D	L	E	E	G	I	Q	T	L	M	125
G	R	L	E	D	G	S	P	R	T	G	Q	I	P	K	Q	T	**Y**	S	K	**F**	D	T	N	S	150
H	N	D	D	A	L	L	K	N	**Y**	G	L	L	**Y**	C	**F**	R	K	D	M	D	K	V	E	T	175
F	L	R	I	V	Q	C	R	S	V	E	G	S	C	G	**F**										191

Fig.2.

v Nde 1

GGATCCTTTTTGTTTAACTTTAAGAAGGAGATATACAT ATG CGT CCG 47
TCT ATC CAC CGT ACC GCT ATC GCT GCT GTT CTG GCT ACC 86
GCT TTC GTT GCT GGT ACC GCT CTG GCA TTC CCG ACC ATC 125
CCG CTG TCT CGT CTG TTC GAC AAC GCT ATG CTG CGT GCT 164
CAC CGT CTG CAC CAG CTG GCT TTC GAC ACC TAC CAG GAA 203
TGG GAA GAA GCT TAC ATC CCG AAA GAA CAG AAA TAC TCT 242
TTC CTG CAG AAC CCG CAG ACC TCT CTG TGC TTC TCT GAA 281
TCT ATC CCG ACC CCG TCT AAC CGT GAA GAA ACC CAG CAG 320
AAA TCT AAC CTG GAA CTG CTG CGT ATC TCT CTG CTG CTG 359
ATC CAG TCT TGG CTG GAA CCG GTT CAG TTC CTG CGT TCT 398
GTT TGG GCT AAC TCT CTG GTT TAC GGT GCT TCT GAC TCT 437
AAC GTT TAC GAC CTG CTG AAA GAC CTG GAA GAA GGT ATC 476
CAG ACC CTG ATG GGT CGT CTG GAA GAC GGT TCT CCG CGT 515
ACC GGT CAG ATC TTC AAA CAG ACC TAC TCT AAA TTC GAC 554
ACC AAC TCT CAC AAC GAC GAC GCT CTG CTG AAA AAC TAC 593
GGT CTG CTG TAC TGC TTC CGT AAA GAC ATG GAC AAA GTT 632
GAA ACC TTC CTG CGT ATC GTT CAG TGC CGT TCT GTT GAA 671

v Xho 1

GGT TCT TGC GGT TTC TAA CTC GAG 695

Fig.4.

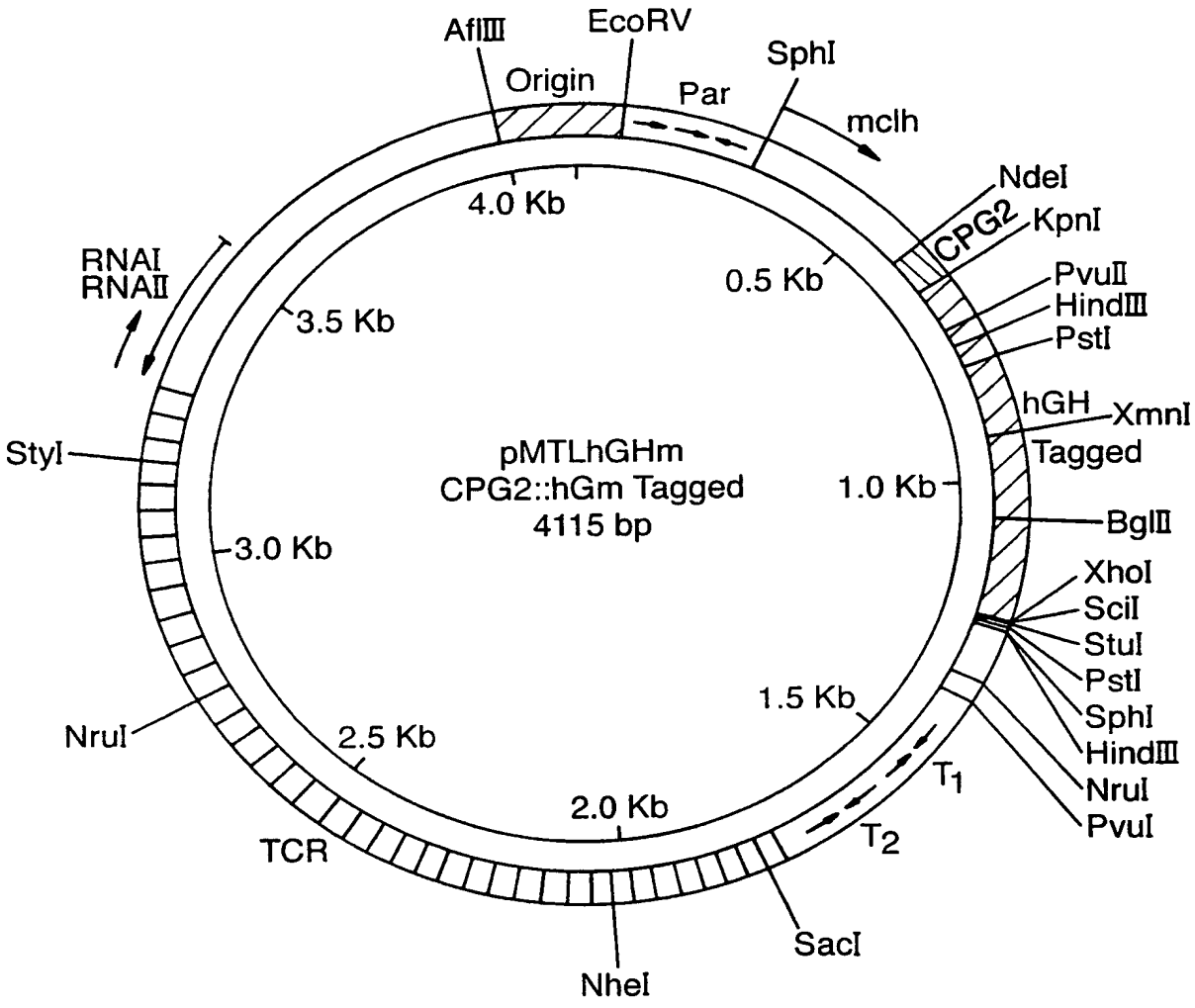

AflIII
EcoRV
SphI
Origin
Par
mclh
NdeI
KpnI
CPG2
PvuII
HindIII
PstI
hGH
XmnI
Tagged
BglII
XhoI
SciI
StuI
PstI
SphI
HindIII
NruI
PvuI
RNAI
RNAII
StyI
NruI
TCR
NheI
SacI
T1
T2
4.0 Kb
0.5 Kb
1.0 Kb
1.5 Kb
2.0 Kb
2.5 Kb
3.0 Kb
3.5 Kb
pMTLhGHm
CPG2::hGm Tagged
4115 bp

Fig.3.

F	P	T	I	P	L	S	R	L	F	D	N	A	M	L	R	A	H	R	L	H	Q	L	A	F	25
D	T	Y	Q	E	**W**	E	E	A	Y	I	P	K	E	Q	K	Y	S	F	L	Q	N	P	Q	T	50
S	L	C	F	S	E	S	I	P	T	P	S	N	R	E	E	T	Q	Q	K	S	N	L	E	L	75
L	R	I	S	L	L	L	I	Q	S	W	L	E	P	V	Q	F	L	R	S	V	**W**	A	N	S	100
L	V	Y	G	A	S	D	S	N	V	Y	D	L	L	K	D	L	E	E	G	I	Q	T	L	M	125
G	R	L	E	D	G	S	P	R	T	G	Q	I	P	K	Q	T	Y	S	K	F	D	T	N	S	150
H	N	D	D	A	L	L	K	N	Y	G	L	L	Y	C	F	R	K	D	M	D	K	V	E	T	175
F	L	R	I	V	Q	C	R	S	V	E	G	S	C	G	F										191

Fig.5.

C	G	N	L	S	T	C	M	L	G	T	Y	T	Q	D	F	N	K	F	H	T	F	P	Q	T	25
A	I	G	V	G	A	P-NH_2																			32

Fig.6.

Y	A	D	A	I	F	T	N	S	Y	R	K	V	L	G	Q	L	S	A	R	K	L	L	Q	D	25
I	M	S	R	Q	Q	G	E	S	N	Q	E	R	G	A	R	R	R	L-NH_2							44

Fig.7A.

Chain A

G I V E Q C C T S I C S L Y Q L E N Y C N 21

Fig.7B.

Chain B

F V N Q H L C G S H L V E A L Y L V C G E R G F F 25

Y T P K T

Fig.8.

A P P R L I C D S R V L Q R Y L L E A K E A E N I 25
T T G C A E H C S L N E N I T V P D T K V N F Y A 50
W K R M E V G Q Q A V E V W Q G L A L L S E A V L 75
R G Q A L L V N S S Q P W E P L Q L H V D K A V S 100
G L R S L T T L L R A L G A Q K E A I S P P D A A 125
S A A P L R T I T A D T F R K L F R V Y S N F L R 150
G K L K L Y T G E A C R T G D 165

Fig.9.

A P T S S S T K K T Q L Q L E H L L L D L Q M I L 25
N G I N N Y K N P K L T R M L T F K F Y M P K K A 50
T E L K H L Q C L E E E L K P L E E V L N L A Q S 75
K N F H L R P R D L I S N I N V I V L E L K G S E 100
T T F M C E Y A D E T A T I V E F L N R W I T F C 125
Q S I I S T L T 133

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Generic Biologicals Limited
(B) STREET: 8 Centre One, Old Sarum Park, Lysander Way
(C) CITY: Salisbury, Wiltshire
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): SP4 6BU
(G) TELEPHONE: (01722) 415026
(H) TELEFAX: (01722) 415028

(ii) TITLE OF INVENTION: Improvements in or Relating to Detection of Molecules in Samples

(iii) NUMBER OF SEQUENCES: 9

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg
1 5 10 15

Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu
20 25 30

Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro
35 40 45

Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg
50 55 60

Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu
65 70 75 80

Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val
85 90 95

Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp

100 105 110

Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu
115 120 125

Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Pro Lys Gln Thr Tyr Ser
130 135 140

Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr
145 150 155 160

Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe
165 170 175

Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
180 185 190

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 695 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGATCCTTTT TGTTTAACTT TAAGAAGGAG ATATACATAT GCGTCCGTCT ATCCACCGTA 60

CCGCTATCGC TGCTGTTCTG GCTACCGCTT TCGTTGCTGG TACCGCTCTG GCATTCCCGA 120

CCATCCCGCT GTCTCGTCTG TTCGACAACG CTATGCTGCG TGCTCACCGT CTGCACCAGC 180

TGGCTTTCGA CACCTACCAG GAATGGGAAG AAGCTTACAT CCCGAAAGAA CAGAAATACT 240

CTTTCCTGCA GAACCCGCAG ACCTCTCTGT GCTTCTCTGA ATCTATCCCG ACCCCGTCTA 300

ACCGTGAAGA AACCCAGCAG AAATCTAACC TGGAACTGCT GCGTATCTCT CTGCTGCTGA 360

TCCAGTCTTG GCTGGAACCG GTTCAGTTCC TGCGTTCTGT TTGGGCTAAC TCTCTGGTTT 420

ACGGTGCTTC TGACTCTAAC GTTTACGACC TGCTGAAAGA CCTGGAAGAA GGTATCCAGA 480

CCCTGATGGG TCGTCTGGAA GACGGTTCTC CGCGTACCGG TCAGATCTTC AAACAGACCT 540

ACTCTAAATT CGACACCAAC TCTCACAACG ACGACGCTCT GCTGAAAAAC TACGGTCTGC 600

TGTACTGCTT CCGTAAAGAC ATGGACAAAG TTGAAACCTT CCTGCGTATC GTTCAGTGCC 660

GTTCTGTTGA AGGTTCTTGC GGTTTCTAAC TCGAG 695

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg
1 5 10 15

Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Trp Glu
20 25 30

Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro
35 40 45

Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg
50 55 60

Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu
65 70 75 80

Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val
85 90 95

Trp Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp
100 105 110

Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu
115 120 125

Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Pro Lys Gln Thr Tyr Ser
130 135 140

Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr
145 150 155 160

Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe
165 170 175

Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
180 185 190

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Cys Gly Asn Leu Ser Thr Cys Met Leu Gly Thr Tyr Thr Gln Asp Phe
1 5 10 15

Asn Lys Phe His Thr Phe Pro Gln Thr Ala Ile Gly Val Gly Ala Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Tyr Ala Asp Ala Ile Phe Thr Asn Ser Tyr Arg Lys Val Leu Gly Gln
1 5 10 15

Leu Ser Ala Arg Lys Leu Leu Gln Asp Ile Met Ser Arg Gln Gln Gly
20 25 30

Glu Ser Asn Gln Glu Arg Gly Ala Arg Arg Arg Leu
35 40

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
1 5 10 15

Glu Asn Tyr Cys Asn
20

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1 5 10 15

Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr
20 25 30

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 165 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Gln Arg Tyr Leu
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
20 25 30

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
65 70 75 80

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
115 120 125

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145 150 155 160

Cys Arg Thr Gly Asp
165

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 133 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu Gln Leu Glu His
1 5 10 15

Leu Leu Leu Asp Leu Gln Met Ile Leu Asn Gly Ile Asn Asn Tyr Lys
20 25 30

Asn Pro Lys Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys
35 40 45

Lys Ala Thr Glu Leu Lys His Leu Gln Cys Leu Glu Glu Glu Leu Lys
50 55 60

Pro Leu Glu Glu Val Leu Asn Leu Ala Gln Ser Lys Asn Phe His Leu
65 70 75 80

Arg Pro Arg Asp Leu Ile Ser Asn Ile Asn Val Ile Val Leu Glu Leu
85 90 95

Lys Gly Ser Glu Thr Thr Phe Met Cys Glu Tyr Ala Asp Glu Thr Ala
100 105 110

Thr Ile Val Glu Phe Leu Asn Arg Trp Ile Thr Phe Cys Gln Ser Ile
115 120 125

Ile Ser Thr Leu Thr
130

INTERNATIONAL SEARCH REPORT

Int ational Application No
PCT/GB 98/03449

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01N33/74 G01N33/58

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and. where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 534 617 A (CUNNINGHAM, BRIAN C. ET AL) 9 July 1996	13-20, 22-24
X	& CHEMICAL ABSTRACTS, vol. 125, no. 15, 7 October 1996 Columbus, Ohio, US; abstract no. 186666, CUNNINGHAM, BRIAN C. ET AL: "Human growth hormone variants having greater affinity for human growth hormone receptor at site 1" See registry nr.180856-17-1 (F10W) see abstract	13-20, 22-24
	--- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 March 1999	06/04/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hoekstra, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int ational Application No
PCT/GB 98/03449

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category '	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 90 04788 A (GENENTECH, INC., USA) 3 May 1990	13-24
X	& CHEMICAL ABSTRACTS, vol. 114, no. 21, 27 May 1991 Columbus, Ohio, US; abstract no. 200473, WELLS, JAMES A. ET AL:: "Method for identification of active domains in polypeptide hormones and other polypeptides using single and multiple amino acid substitutions" See registry nr.130911-61-4 (F10W) and 130911-63-6 (F97W) see abstract	13-24
X	WO 92 09690 A (GENENTECH, INC., USA) 11 June 1992 See (18) and (20): F10W see claim 48	13-20, 22-24
A	WO 94 10200 A (UPJOHN CO ;LEHRMAN S RUSS (US)) 11 May 1994 cited in the application see the whole document	1-26
A	PLATER, MICHAEL L. ET AL: "Effects of site-directed mutations on the chaperone-like activity of.alpha.B-crystallin" J. BIOL. CHEM. (1996), 271(45), 28558-28566 CODEN: JBCHA3;ISSN: 0021-9258, XP002095857	1-26

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB 98/03449

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 25,26
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
Claims 25 and 26 do not refer to any technical features and hence do not comply with the prescribed requirements of article 6 and rule 6.3(a) PCT to such an extend that a meaningful search is not possible.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 98/03449

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5534617 A	09-07-1996	US 5750373 A	12-05-1998
		US 5854026 A	29-12-1998
		US 5580723 A	03-12-1996
		US 5766854 A	16-06-1998
		US 5834250 A	10-11-1998
		US 5846765 A	08-12-1998
		US 5821047 A	13-10-1998
		US 5834598 A	10-11-1998
		AT 164395 T	15-04-1998
		CA 2095633 A	04-06-1992
		DE 69129154 D	30-04-1998
		DE 69129154 T	20-08-1998
		EP 0564531 A	13-10-1993
		ES 2113940 T	16-05-1998
		GR 3026468 T	30-06-1998
		WO 9209690 A	11-06-1992
		US 5780279 A	14-07-1998
		JP 7503600 T	20-04-1995
		US 5688666 A	18-11-1997
		EP 0397834 A	22-11-1990
		JP 4502454 T	07-05-1992
		WO 9004788 A	03-05-1990
		CA 2001774 A	28-04-1990
WO 9004788 A		NONE	
WO 9209690 A		NONE	
WO 9410200 A		NONE	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/03449

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01N33/74 G01N33/58

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 534 617 A (CUNNINGHAM, BRIAN C. ET AL) 9 July 1996	13-20, 22-24
X	& CHEMICAL ABSTRACTS, vol. 125, no. 15, 7 October 1996 Columbus, Ohio, US; abstract no. 186666, CUNNINGHAM, BRIAN C. ET AL: "Human growth hormone variants having greater affinity for human growth hormone receptor at site 1" See registry nr.180856-17-1 (F10W) see abstract	13-20, 22-24
	--- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 March 1999	06/04/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hoekstra, S

2

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 90 04788 A (GENENTECH, INC., USA) 3 May 1990	13-24
X	& CHEMICAL ABSTRACTS, vol. 114, no. 21, 27 May 1991 Columbus, Ohio, US; abstract no. 200473, WELLS, JAMES A. ET AL:: "Method for identification of active domains in polypeptide hormones and other polypeptides using single and multiple amino acid substitutions" See registry nr.130911-61-4 (F10W) and 130911-63-6 (F97W) see abstract	13-24
X	WO 92 09690 A (GENENTECH, INC., USA) 11 June 1992 See (18) and (20): F10W see claim 48	13-20, 22-24
A	WO 94 10200 A (UPJOHN CO ;LEHRMAN S RUSS (US)) 11 May 1994 cited in the application see the whole document	1-26
A	PLATER, MICHAEL L. ET AL: "Effects of site-directed mutations on the chaperone-like activity of.alpha.B-crystallin" J. BIOL. CHEM. (1996), 271(45), 28558-28566 CODEN: JBCHA3;ISSN: 0021-9258, XP002095857	1-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 98/03449

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5534617 A	09-07-1996	US 5750373 A	12-05-1998
		US 5854026 A	29-12-1998
		US 5580723 A	03-12-1996
		US 5766854 A	16-06-1998
		US 5834250 A	10-11-1998
		US 5846765 A	08-12-1998
		US 5821047 A	13-10-1998
		US 5834598 A	10-11-1998
		AT 164395 T	15-04-1998
		CA 2095633 A	04-06-1992
		DE 69129154 D	30-04-1998
		DE 69129154 T	20-08-1998
		EP 0564531 A	13-10-1993
		ES 2113940 T	16-05-1998
		GR 3026468 T	30-06-1998
		WO 9209690 A	11-06-1992
		US 5780279 A	14-07-1998
		JP 7503600 T	20-04-1995
		US 5688666 A	18-11-1997
		EP 0397834 A	22-11-1990
		JP 4502454 T	07-05-1992
		WO 9004788 A	03-05-1990
		CA 2001774 A	28-04-1990
WO 9004788 A		NONE	
WO 9209690 A		NONE	
WO 9410200 A		NONE	

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
G01N 33/74, 33/58

A1

(11) International Publication Number: WO 99/26069

(43) International Publication Date: 27 May 1999 (27.05.99)

(21) International Application Number: PCT/GB98/03449

(22) International Filing Date: 16 November 1998 (16.11.98)

(30) Priority Data:
9723955.2 14 November 1997 (14.11.97) GB

(71) Applicant *(for all designated States except US)*: GENERIC BIOLOGICALS LIMITED [GB/GB]; 8 Centre One, Old Sarum Park, Lysander Way, Salisbury, Wiltshire SP4 6BU (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MURPHY, Jonathan, Paul [GB/GB]; 37 Bishops Orchard, East Hagbourne, Oxfordshire OX11 9JS (GB). ATKINSON, Anthony [GB/GB]; Twingley, Mill Corner, Winterbourne Gunner, Salisbury, Wiltshire SP4 6JJ (GB).

(74) Agent: KEITH W. NASH & CO.; 90–92 Regent Street, Cambridge CB2 1DP (GB).

(81) Designated States: JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMPROVEMENTS IN OR RELATING TO DETECTION OF MOLECULES IN SAMPLES

```
F P T I P L S R L F D N A M L R A H R L H Q L A F 25
D T Y Q E W E E A Y I P K E Q K Y S F L Q N P Q T 50
S L C F S E S I P T P S N R E E T Q Q K S N L E L 75
L R I S L L L I Q S W L E P V Q F L R S V W A N S 100
L V Y G A S D S N V Y D L L K D L E E G I Q T L M 125
G R L E D G S P R T G Q I P K Q T Y S K F D T N S 150
H N D D A L L K N Y G L L Y C F R K D M D K V E T 175
F L R I V Q C R S V E G S C G F                   191
```

(57) Abstract

Disclosed is a method of detecting the presence in a sample of a polypeptide exogenously administered to a mammalian subject from whom the sample is obtained, and distinguishing between such an exogenously administered polypeptide and a naturally–occurring endogenous polypeptide present in the sample; the method comprising obtaining a sample from the subject; and subjecting the sample to analysis of fluorescence at a suitable wavelength; wherein the exogenously administered polypeptide is tagged with a greater or lesser amount of fluorescence activity, relative to the untagged endogenous polypeptide, at the wavelength(s) analysed.

0 9 / ...

For receiving Office use only

International Application No.

International Filing Date

Name of receiving Office and "PCT International Application"

Applicant's or agent's file reference (if desired) (12 characters maximum) M L/C320.01/G

PCT

RE...

...ests that the present ...ation be processed ...ent Cooperation Treaty.

The un... int... acc...

...LE OF INVENTION

...nts in or Relating to Detection of Molecules in Samples

... APPLICANT

...address: *(Family name followed by given name; for a legal entity, full official ...ion. The address must include postal code and name of country. The country of the ...s indicated in this Box is the applicant's State (that is, country) of residence if no State ...idence is indicated below.)*

Generic Biologicals Limited
8 Centre One, Old Sarum Park
Lysander Way, Salisbury
Wiltshire SP4 6BU
United Kingdom

[ ] This person is also inventor.

Telephone No. (01722) 415026

Facsimile No. (01722) 415028

Teleprinter No.

State *(that is, country)* of nationality: GB

State *(that is, country)* of residence: GB

This person is applicant for the purposes of: [ ] all designated States [X] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

MURPHY, Jonathan Paul
37 Bishops Orchard
East Hagbourne
Oxfordshire OX11 9JS
United Kingdom

This person is:

[ ] applicant only

[X] applicant and inventor

[ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality: GB

State *(that is, country)* of residence: GB

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [X] the United States of America only [ ] the States indicated in the Supplemental Box

[X] Further applicants and/or (further) inventors are indicated on a continuation sheet.

Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: [X] agent [ ] common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

Keith W Nash & Co
90-92 Regent Street
Cambridge CB2 1DP
United Kingdom

Telephone No. (01223) 355477

Facsimile No. (01223) 324353

Teleprinter No.

[ ] **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

See Notes to the request form

Continuation of Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

If none of the following sub-boxes is used, this sheet should not be included in the request.

Name and address: *(Family name followed by given name: for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

ATKINSON, Anthony

Twingley, Mill Corner
Winterbourne Gunner, Salisbury
Wiltshire SP4 6JJ
United Kingdom

This person is:
- [ ] applicant only
- [X] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality: GB

State *(that is, country)* of residence: GB

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [X] the United States of America only [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name: for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name: for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name: for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

[ ] Further applicants and/or (further) inventors are indicated on another continuation sheet.

Box No.V DESIGNATION OF STATES

The following designations are hereby made under Rule 4.9(a) *(mark the applicable check-boxes: at least one must be marked)*:

Regional Patent

- ☐ AP **ARIPO Patent:** GH Ghana, GM Gambia, KE Kenya, LS Lesotho, MW Malawi, SD Sudan, SZ Swaziland, UG Uganda, ZW Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT
- ☐ EA **Eurasian Patent:** AM Armenia, AZ Azerbaijan, BY Belarus, KG Kyrgyzstan, KZ Kazakhstan, MD Republic of Moldova, RU Russian Federation, TJ Tajikistan, TM Turkmenistan, and any other State which is a Contracting State of the Eurasian Patent Convention and of the PCT
- ☒ EP **European Patent:** AT Austria, BE Belgium, CH and LI Switzerland and Liechtenstein, CY Cyprus, DE Germany, DK Denmark, ES Spain, FI Finland, FR France, GB United Kingdom, GR Greece, IE Ireland, IT Italy, LU Luxembourg, MC Monaco, NL Netherlands, PT Portugal, SE Sweden, and any other State which is a Contracting State of the European Patent Convention and of the PCT
- ☐ OA **OAPI Patent:** BF Burkina Faso, BJ Benin, CF Central African Republic, CG Congo, CI Côte d'Ivoire, CM Cameroon, GA Gabon, GN Guinea, ML Mali, MR Mauritania, NE Niger, SN Senegal, TD Chad, TG Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT *(if other kind of protection or treatment desired, specify on dotted line)* .....

National Patent *(if other kind of protection or treatment desired, specify on dotted line)*:

- ☐ AL Albania .....
- ☐ AM Armenia .....
- ☐ AT Austria .....
- ☐ AU Australia .....
- ☐ AZ Azerbaijan
- ☐ BA Bosnia and Herzegovina .....
- ☐ BB Barbados
- ☐ BG Bulgaria .....
- ☐ BR Brazil .....
- ☐ BY Belarus .....
- ☐ CA Canada
- ☐ CH and LI Switzerland and Liechtenstein
- ☐ CN China .....
- ☐ CU Cuba .....
- ☐ CZ Czech Republic .....
- ☐ DE Germany .....
- ☐ DK Denmark .....
- ☐ EE Estonia .....
- ☐ ES Spain .....
- ☐ FI Finland .....
- ☐ GB United Kingdom
- ☐ GE Georgia .....
- ☐ GH Ghana .....
- ☐ GM Gambia
- ☐ GW Guinea-Bissau .....
- ☐ HR Croatia
- ☐ HU Hungary .....
- ☐ ID Indonesia
- ☐ IL Israel .....
- ☐ IS Iceland
- ☒ JP Japan .....
- ☐ KE Kenya .....
- ☐ KG Kyrgyzstan .....
- ☐ KP Democratic People's Republic of Korea .....
- ☐ KR Republic of Korea .....
- ☐ KZ Kazakhstan .....
- ☐ LC Saint Lucia
- ☐ LK Sri Lanka
- ☐ LR Liberia
- ☐ LS Lesotho .....
- ☐ LT Lithuania
- ☐ LU Luxembourg
- ☐ LV Latvia
- ☐ MD Republic of Moldova .....
- ☐ MG Madagascar .....
- ☐ MK The former Yugoslav Republic of Macedonia .....
- ☐ MN Mongolia
- ☐ MW Malawi .....
- ☐ MX Mexico .....
- ☐ NO Norway
- ☐ NZ New Zealand .....
- ☐ PL Poland .....
- ☐ PT Portugal .....
- ☐ RO Romania
- ☐ RU Russian Federation .....
- ☐ SD Sudan
- ☐ SE Sweden
- ☐ SG Singapore
- ☐ SI Slovenia .....
- ☐ SK Slovakia .....
- ☐ SL Sierra Leone
- ☐ TJ Tajikistan .....
- ☐ TM Turkmenistan .....
- ☐ TR Turkey .....
- ☐ TT Trinidad and Tobago .....
- ☐ UA Ukraine .....
- ☐ UG Uganda .....
- ☒ US United States of America .....
- ☐ UZ Uzbekistan .....
- ☐ VN Viet Nam .....
- ☐ YU Yugoslavia .....
- ☐ ZW Zimbabwe .....

Check-boxes reserved for designating States (for the purposes of a national patent) which have become party to the PCT after issuance of this sheet:

- ☐ .....
- ☐ .....

Precautionary Designation Statement: In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except any designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. *(Confirmation of a designation consists of the filing of a notice specifying that designation and the payment of the designation and confirmation fees. Confirmation must reach the receiving Office within the 15-month time limit.)*

See Notes to the request form

Box No. VI PRIORITY CLAIM ☐ Further priority claims are indicated in the Supplemental Box.

Filing date of earlier application (day/month/year)	Number of earlier application	Where earlier application is: national application: country	regional application:* regional Office	international application: receiving Office
item (1) 14/11/1997	9723955.2	GB		
item (2)				
item (3)				

☒ The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application was filed with the Office which for the purposes of the present international application is the receiving Office)* identified above as item(s): (1)

* *Where the earlier application is an ARIPO application, it is mandatory to indicate in the Supplemental Box at least one country party to the Paris Convention for the Protection of Industrial Property for which that earlier application was filed (Rule 4.10(b)(ii)). See Supplemental Box.*

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) *(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used)*: ISA /

Request to use results of earlier search; reference to that search *(if an earlier search has been carried out by or requested from the International Searching Authority)*: Date *(day/month/year)* Number Country *(or regional Office)*

Box No. VIII CHECK LIST; LANGUAGE OF FILING

This international application contains the following **number of sheets**:

request	:	4
description (excluding sequence listing part)	:	19
claims	:	3
abstract	:	1
drawings	:	5
sequence listing part of description	:	6
Total number of sheets	:	38

This international application is **accompanied by** the item(s) marked below:

1. ☒ fee calculation sheet
2. ☒ separate signed power of attorney
3. ☐ copy of general power of attorney; reference number, if any:
4. ☐ statement explaining lack of signature
5. ☐ priority document(s) identified in Box No. VI as item(s):
6. ☐ translation of international application into (language):
7. ☐ separate indications concerning deposited microorganism or other biological material
8. ☒ nucleotide and/or amino acid sequence listing in computer readable form
9. ☐ other *(specify)*:

Figure of the drawings which should accompany the abstract: 3

Language of filing of the international application: ENGLISH

Box No. IX SIGNATURE OF APPLICANT OR AGENT

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).

Keith W. Nash & Co.

Keith W Nash & Co (Agents)

For receiving Office use only

1. Date of actual receipt of the purported international application:	2. Drawings: ☐ received: ☐ not received:
3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:	
4. Date of timely receipt of the required corrections under PCT Article 11(2):	
5. International Searching Authority (if two or more are competent): ISA /	6. ☐ Transmittal of search copy delayed until search fee is paid.

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

PATENT COOPERATION TREATY

09/554451

PCT

From the INTERNATIONAL BUREAU

To:

KEITH W NASH & CO
90-92 Regent Street
Cambridge CB2 1DP
ROYAUME-UNI

NOTIFICATION CONCERNING SUBMISSION OR TRANSMITTAL OF PRIORITY DOCUMENT

(PCT Administrative Instructions, Section 411)

Date of mailing (day/month/year) 05 January 1999 (05.01.99)	
Applicant's or agent's file reference M L/C320.01/G	**IMPORTANT NOTIFICATION**
International application No. PCT/GB98/03449	International filing date (day/month/year) 16 November 1998 (16.11.98)
International publication date (day/month/year) Not yet published	Priority date (day/month/year) 14 November 1997 (14.11.97)
Applicant GENERIC BIOLOGICALS LIMITED et al	

1. The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).

2. This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.

3. An **asterisk(*)** appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

4. The **letters "NR"** appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, **the attention of the applicant is directed** to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

Priority date	Priority application No.	Country or regional Office or PCT receiving Office	Date of receipt of priority document
14 Nove 1997 (14.11.97)	9723955.2	GB	18 Dece 1998 (18.12.98)

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No. (41-22) 740.14.35

Authorized officer

Juan Cruz

Telephone No. (41-22) 338.83.38

PATENT COOPERATION TREATY

09/554451

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
MJL/C320.01/G		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/GB98/03449	16/11/1998	14/11/1997
International Patent Classification (IPC) or national classification and IPC		
G01N33/74		
Applicant		
GENERIC BIOLOGICALS LIMITED et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☒ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
17/05/1999	20. 12. 99
Name and mailing address of the international preliminary examining authority:	Authorized officer
European Patent Office - P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Tx: 31 651 epo nl Fax: +31 70 340 - 3016	Hoekstra, S Telephone No. +31 70 340 2847

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/GB98/03449

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-19 as originally filed

Claims, No.:

1-24 as originally filed

Drawings, sheets:

1/5-5/5 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application.

☒ claims Nos. 25 and 26.

because:

☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):

☒ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 25 and 26 are so unclear that no meaningful opinion could be formed (*specify*):

see separate sheet

☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

☒ no international search report has been established for the said claims Nos. 25 and 26.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims 1-12
	No:	Claims 13-24
Inventive step (IS)	Yes:	Claims 1-12
	No:	Claims 13-24
Industrial applicability (IA)	Yes:	Claims 1-24
	No:	Claims

2. Citations and explanations

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

Re Item V

Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. The documents cited in the search report do not disclose or suggest to differentiate in a body sample between exogenously administered and endogenous polypeptides by replacing amino acids for fluorescing amino acids in the backbone of the polypeptide. The methods of claims 8 to 12, referring to all essential features of the invention are hence considered to meet the requirements of Article 33 PCT. Provisional to the incorporation of all essential features of the invention also claims 1-7 can be considered to meet the requirements of Article 33(2) and 33(3) PCT.

2. The present application does, however, not meet the requirements of Article 33(2) because the subject-matter of claims 13-24 encompass matter known in the art.

   a. WO9209690 disclosed hGH mutants [(18) and (20) of claim 48] having a Phe residue at position 10 replaced by Trp.

   b. CA114(21), No. 200473 (WO9004788) disclosed registry nr. 130911-63-7: a hGH having Phe residue at position 97 replaced by Trp and registry nr. 130911-61-4: a hGH having Phe at position 10 replaced by Trp.

   c. CA125(15), No.186666 (US5534617) disclosed registry nr. 180856-17-1: a hGH having a Phe residue at position 10 replaced by Trp.

3. These hGH mutants of the prior art are intended for use in compositions for administering to a mammalian subject. The separate teachings of the above three documents taken alone destroy the novelty of claims 13-24.

4. In view of the host of hGH mutants available in the state of the art it is considered that any generalised formulation of hGH mutants or of any of the other mentioned (but not claimed) proteins have to be considered obvious.

5. Mutants of hGH per se are obvious solutions to the technical problem of providing alternative proteins having hGH activity and any further desired properties.

6. Compositions comprising specific hGH mutants which are a solution to the problem of providing tagged hGH molecules are considered, provided novelty is established, to involve inventive skill. However, no claims limited to such hGH mutants only is present in the set of claims forming a basis for this opinion.

7. In assessing novelty for products (here compositions characterised by hGH mutants) the intended use of the product is irrelevant.

Re Item VII

Certain defects in the international application

8. The following defects in the form or contents of the international application have been noted:

Claims 1-7 and 13 encompass in their broadest outline also the use of polypeptides tagged with conventional fluorophores other than amino acids covalently attached other than via peptide bonds to the backbone of the polypeptide. In view of the concept of the present invention it is considered that only tags consisting of fluorescing amino acids replacing amino acids in the polypeptide backbone attribute to the solution presented in this application to the underlying technical problem of differentiating exogenously administered protein from endogenous ones.As such the features of the tags being amino acids replacing natural amino acid residues is an essential feature of the invention and must hence be a limiting feature of the claims (Article 6 and Rule 6.3(a) PCT).

PATENT COOPERATION TREATY 09 / 554513

PCT

REC'D 23 DEC 1999
WIPO PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
MJL/C320.01/G		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/GB98/03449	16/11/1998	14/11/1997
International Patent Classification (IPC) or national classification and IPC G01N33/74		
Applicant GENERIC BIOLOGICALS LIMITED et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☒ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
17/05/1999	20. 12. 99
Name and mailing address of the international preliminary examining authority: European Patent Office - P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Tx: 31 651 epo nl Fax: +31 70 340 - 3016	Authorized officer Hoekstra, S Telephone No. +31 70 340 2847

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/GB98/03449

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-19 as originally filed

Claims, No.:

1-24 as originally filed

Drawings, sheets:

1/5-5/5 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application.

☒ claims Nos. 25 and 26.

because:

☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):

☒ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 25 and 26 are so unclear that no meaningful opinion could be formed (*specify*):

see separate sheet

☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

☒ no international search report has been established for the said claims Nos. 25 and 26.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-12
	No:	Claims	13-24
Inventive step (IS)	Yes:	Claims	1-12
	No:	Claims	13-24
Industrial applicability (IA)	Yes:	Claims	1-24
	No:	Claims	

2. Citations and explanations

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

Re Item V

Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. The documents cited in the search report do not disclose or suggest to differentiate in a body sample between exogenously administered and endogenous polypeptides by replacing amino acids for fluorescing amino acids in the backbone of the polypeptide. The methods of claims 8 to 12, referring to all essential features of the invention are hence considered to meet the requirements of Article 33 PCT. Provisional to the incorporation of all essential features of the invention also claims 1-7 can be considered to meet the requirements of Article 33(2) and 33(3) PCT.

2. The present application does, however, not meet the requirements of Article 33(2) because the subject-matter of claims 13-24 encompass matter known in the art.

   a. WO9209690 disclosed hGH mutants [(18) and (20) of claim 48] having a Phe residue at position 10 replaced by Trp.

   b. CA114(21), No. 200473 (WO9004788) disclosed registry nr. 130911-63-7: a hGH having Phe residue at position 97 replaced by Trp and registry nr. 130911-61-4: a hGH having Phe at position 10 replaced by Trp.

   c. CA125(15), No.186666 (US5534617) disclosed registry nr. 180856-17-1: a hGH having a Phe residue at position 10 replaced by Trp.

3. These hGH mutants of the prior art are intended for use in compositions for administering to a mammalian subject. The separate teachings of the above three documents taken alone destroy the novelty of claims 13-24.

4. In view of the host of hGH mutants available in the state of the art it is considered that any generalised formulation of hGH mutants or of any of the other mentioned (but not claimed) proteins have to be considered obvious.

5. Mutants of hGH per se are obvious solutions to the technical problem of providing alternative proteins having hGH activity and any further desired properties.

6. Compositions comprising specific hGH mutants which are a solution to the problem of providing tagged hGH molecules are considered, provided novelty is established, to involve inventive skill. However, no claims limited to such hGH mutants only is present in the set of claims forming a basis for this opinion.

7. In assessing novelty for products (here compositions characterised by hGH mutants) the intended use of the product is irrelevant.

Re Item VII

Certain defects in the international application

8. The following defects in the form or contents of the international application have been noted:

Claims 1-7 and 13 encompass in their broadest outline also the use of polypeptides tagged with conventional fluorophores other than amino acids covalently attached other than via peptide bonds to the backbone of the polypeptide. In view of the concept of the present invention it is considered that only tags consisting of fluorescing amino acids replacing amino acids in the polypeptide backbone attribute to the solution presented in this application to the underlying technical problem of differentiating exogenously administered protein from endogenous ones.As such the features of the tags being amino acids replacing natural amino acid residues is an essential feature of the invention and must hence be a limiting feature of the claims (Article 6 and Rule 6.3(a) PCT).

T.J

PATENT COOPERATION TREATY

09/554451

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION** see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
M L/C320.01/G		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/GB 98/ 03449	16/11/1998	14/11/1997
Applicant		
GENERIC BIOLOGICALS LIMITED et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 4 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language,** the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

[ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

[ ] contained in the international application in written form.

[ ] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readble form.

[ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [X] **Certain claims were found unsearchable** (See Box I).

3. [ ] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. 3

[X] as suggested by the applicant.

[ ] None of the figures.

[ ] because the applicant failed to suggest a figure.

[ ] because this figure better characterizes the invention.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB 98/ 03449

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [X] Claims Nos.: 25,26
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
Claims 25 and 26 do not refer to any technical features and hence do not comply with the prescribed requirements of article 6 and rule 6.3(a) PCT to such an extend that a meaningful search is not possible.

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/03449

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G01N33/74 G01N33/58

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G01N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 534 617 A (CUNNINGHAM, BRIAN C. ET AL) 9 July 1996	13-20, 22-24
X	& CHEMICAL ABSTRACTS, vol. 125, no. 15, 7 October 1996 Columbus, Ohio, US; abstract no. 186666, CUNNINGHAM, BRIAN C. ET AL: "Human growth hormone variants having greater affinity for human growth hormone receptor at site 1" See registry nr.180856-17-1 (F10W) see abstract	13-20, 22-24
	--- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 March 1999	06/04/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hoekstra, S

2

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 90 04788 A (GENENTECH, INC., USA) 3 May 1990	13-24
X	& CHEMICAL ABSTRACTS, vol. 114, no. 21, 27 May 1991 Columbus, Ohio, US; abstract no. 200473, WELLS, JAMES A. ET AL:: "Method for identification of active domains in polypeptide hormones and other polypeptides using single and multiple amino acid substitutions" See registry nr.130911-61-4 (F10W) and 130911-63-6 (F97W) see abstract	13-24
X	WO 92 09690 A (GENENTECH, INC., USA) 11 June 1992 See (18) and (20): F10W see claim 48	13-20, 22-24
A	WO 94 10200 A (UPJOHN CO ;LEHRMAN S RUSS (US)) 11 May 1994 cited in the application see the whole document	1-26
A	PLATER, MICHAEL L. ET AL: "Effects of site-directed mutations on the chaperone-like activity of.alpha.B-crystallin" J. BIOL. CHEM. (1996), 271(45), 28558-28566 CODEN: JBCHA3;ISSN: 0021-9258, XP002095857	1-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 98/03449

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5534617 A	09-07-1996	US 5750373 A	12-05-1998
		US 5854026 A	29-12-1998
		US 5580723 A	03-12-1996
		US 5766854 A	16-06-1998
		US 5834250 A	10-11-1998
		US 5846765 A	08-12-1998
		US 5821047 A	13-10-1998
		US 5834598 A	10-11-1998
		AT 164395 T	15-04-1998
		CA 2095633 A	04-06-1992
		DE 69129154 D	30-04-1998
		DE 69129154 T	20-08-1998
		EP 0564531 A	13-10-1993
		ES 2113940 T	16-05-1998
		GR 3026468 T	30-06-1998
		WO 9209690 A	11-06-1992
		US 5780279 A	14-07-1998
		JP 7503600 T	20-04-1995
		US 5688666 A	18-11-1997
		EP 0397834 A	22-11-1990
		JP 4502454 T	07-05-1992
		WO 9004788 A	03-05-1990
		CA 2001774 A	28-04-1990
WO 9004788 A		NONE	
WO 9209690 A		NONE	
WO 9410200 A		NONE	

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 271, No. 45, Issue of November 8, pp. 28558–28566, 1996
Printed in U.S.A.

Effects of Site-directed Mutations on the Chaperone-like Activity of αB-Crystallin*

PD 1996 P. 28558-566 = (9)

(Received for publication, May 30, 1996, and in revised form, August 12, 1996)

Michael L. Plater‡, Derek Goode, and M. James C. Crabbe§

From the Wolfson Laboratory, School of Animal and Microbial Sciences, The University of Reading, Whiteknights, Reading, Berkshire RG6 6AJ, United Kingdom

Recombinant αB-crystallin has been shown to exhibit chaperone-like activity, suppressing the thermal aggregation of γ-crystallin and aggregation of the reduced insulin B chain conferring thermotolerance to *Escherichia coli* BL21(DE3) cells. Mutations were made in three specific areas of the αB-crystallin, the N terminus D2G, the conserved phenylalanine-rich region, F24R, F27R, F27A, and the two C-terminal lysines K174L/K175L, K174G/K175G. Biophysical characterization of the mutant αB-crystallins using far-UV CD revealed no change in secondary structural elements. Tryptophan fluorescence demonstrated global structural changes. Heat stability of the mutant αB-crystallins was not significantly affected as indicated by tryptophan fluorescence of heat-treated proteins.

Mutations within the phenylalanine-rich region abolish the chaperone-like activity as measured by both *in vivo* and *in vitro* assays. Proteins with mutations at the C terminus demonstrated no significant chaperone-like activity, failing to confer thermotolerance on *E. coli* and demonstrating no significant inhibition of protein aggregation in either γ-crystallin or reduced insulin B chain assays. The N-terminal mutation D2G demonstrated a significant reduction in efficiency of the chaperone-like activity although some thermotolerance was conferred in the *E. coli* assay. *In vitro* assays showed that complete inhibition of aggregation was only achieved at 10-fold higher concentrations of D2G than that required by the native αB-crystallin.

Consistent changes in the chaperone-like activity of the site-directed mutants were demonstrated by the three assays. The results suggested that both charge-charge and hydrophobic interactions are important in protein binding by αB-crystallin and that the conserved RLFDQFF region is vital for chaperone-like activity.

The eye lens contains high concentrations of soluble proteins, the crystallins. They fall into two classes, the α-crystallin family and the β/γ-superfamily (1). There is differential expression of the crystallins during lens development (2) which leads to different mixtures of crystallins along the visual axis. Properties of individual crystallins may be important in maintaining short range order, and thus transparency, in the lens (3). As there is no protein turnover in the majority of the lens tissue, the crystallins can survive as long as the individual. Crystallin unfolding and aggregation, caused initially by post-translational modification (4, 5) or oxidative damage (6) to these long-lived proteins, play an aetiological role in the development of cataract, the largest cause of blindness in the world.

α-Crystallins differ from β- and γ-crystallins in containing both sheet and helical structure and have been found in many extralenticular tissues including brain, spinal cord, and heart (7). Overexpression of αB-crystallin may be induced in mammalian cell lines by both heat shock (8) and osmotic stress (9) and by expression of oncogene proteins such as c-Ha-Ras (10). Expression of αB-crystallin in glial cell culture has been shown to confer increased thermotolerance and adherence and to cause an increase in cytoskelatal fibers (11). Overexpression of αB-crystallin has been observed in a large number of severe neurological disorders including Creutzfeldt-Jakob disease (12), and αB-crystallin serves as an immunodominant myelin antigen to human T cells when expressed at the elevated levels found in active multiple sclerosis lesions (13).

Chemical modification of α- and γ-crystallins by cyanate and by glucose 6-phosphate causes the proteins to alter their tertiary structure, but retain native secondary structure (14). These initial changes correspond to those found in human cataract, where glycation is an important aetiological factor (15).

Many monomeric and oligomeric proteins lack the inherent ability to correctly assemble into biologically functional molecules (16). For the correct post-translational assembly of these polypeptides into their "correct" structures, a ubiquitous class of conserved proteins, termed "chaperones" or, for bacterial proteins, "chaperonins" is thought to be involved. This class of proteins includes the *Escherichia coli* GroEL, SecB, DnaK, and DnaJ gene products, and heat shock proteins (hsp60, -70, -90) in eucaryotic cells. As well as a general class of chaperone proteins, there appears to be a class of specific chaperones that regulate the folding of a single protein (17). Recent evidence supports the idea that GroEL and SecB can act as "unfoldases," binding and unfolding stressed and aberrantly folded proteins and restoring the native structure (18).

Lens α-crystallin, which has some sequence homology to heat shock proteins, was suggested to be a chaperone for correct folding of γ-crystallin in the lens (19), hence maintaining lens transparency. α-Crystallin has been shown to act as a chaperone-like protein, *in vitro* sequestering unfolded protein, and inhibiting its subsequent aggregation and insolubilization (19–21). However, αB-crystallin has been shown to differ from true chaperones in that subsequent release of the bound protein and restoration of native structure has not been observed (22, 23). This may be because αB requires an as yet unidentified cofactor, analogous to GroES, present *in vivo* but not in the *in vitro* experiments or because αB-crystallin belongs to a more simple class of shock proteins which bind unfolded proteins in a chaperone-like manner and prevent aggregation but do not release

* This work was supported by the Wellcome Trust and the Royal National Institute for the Blind. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a studentship from the Medical Research Council.

§ To whom correspondence should be addressed: Wolfson Laboratory, School of Animal and Microbial Sciences, The University of Reading, P. O. Box 228, Whiteknights, Reading, Berkshire RG6 6AJ, UK. Tel. and Fax: 44-01734-318894.

A ATGGACATCGCCATCCACCACCCCTGGATCCGGCGCCCCTTCTTCCC
GCG
CTTCCACTCCCCAAGCCGCCTCTTCGACCAGTTCTTCGGAGAGCACCTG
CGC GGC
CGC
TTGGAGTCTGACCTCTTCTCAACAGCCACTTCCCTGAGCCCCTTCTAC

CTTCGGCCACCCTCCTTCCTGCGGGCACCCAGCTGGATTGACACCGG

ACTCTCAGAGATGCGTTTGGAGAAGGACAGATTCTCTGTGAATCTGG

ACGTGAAGCACTTCTCTCCGGAGGAACTCAAAGTCAAGGTTCTGGGG

GACGTGATTGAGGTCCACGGCAAGCACGAAGAACGCCAGGACGAACA

TGGCTTCATCTCCAGGGAGTTCCACAGGAAGTACCGGATCCCAGCCG

ATGTGGATCCTCTCACCATCACTTCATCCCTGTCATCTGATGGAGTCC

TCACTGTGAATGGACCAAGGAAACAGGTGTCTGGCCCTGAGCGCACC

ATTCCCATCACCCGTGAAGAGAGAAGCCTGCTGTCGCCGCAGCCCCT

AAGAAGTAG.
CTGCTG
GGGGGA

B
MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFSTATSLSPFYLRPPS
G R A
R
FLRAPSWIDTGLSEMRLEKDRFSVNLDKHFSPEELKVKVLGDVIEVHGK

HEERDQDEHGFISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQVS

GPERTIPITREEKPAVAAAPKK
LL
GG

FIG. 1. **DNA sequence (*A*) and amino acid sequence (*B*) of the murine αB-crystallin cDNA and protein, showing loci of the site-directed mutations.** Codons/residues shown in *bold* in the native sequence indicate the position of the substitutions. The substituted sequence for each mutant is given below each site in *italics*. The conserved phenylalanine-rich region thought to be important in chaperone-like activity/aggregate formation in α-crystallin is shown as *underlined italicized* residues in *B*.

them. α-Crystallin is therefore said to have a chaperone-like activity rather than a chaperone activity.

αB-Crystallin does prevent aggregation, however, and since cataract formation involves modifications to crystallins followed by protein unfolding and aggregation events, it has been postulated that α-crystallin chaperone-like function is necessary to maintain lens transparency. Lyophilization of γ-crystallin in the presence of α-crystallin did not alter the structure of the γ-crystallin molecule (24). Evidence for an *in vivo* chaperone-like sequestering function for α-crystallin in the lens has recently been obtained by electron microscopy of immunolabeled α/γ- and α/β-crystallin complexes in bovine lenses (25). α-Crystallin is therefore now of considerable interest both as a lens protein involved in cataractogenesis and as a general mammalian shock protein with a possible role in other disorders.

The small size of the α-crystallin monomers (175 residues) make it a very useful model for determination of specific residues involved in protein binding. Modification of the C-terminal region of αB-crystallin has been shown to inhibit the *in vitro* chaperone-like activity (20). The C terminus contains a number of lysine residues which may be glycated in the aging lens. Protein binding may involve hydrophobic residues, and we have found a considerable sequence homology between a conserved phenylalanine-rich region RLFDQFF in αA- and B-crystallins and a number of heat-shock proteins (26).

We therefore decided to use site-directed mutagenesis to

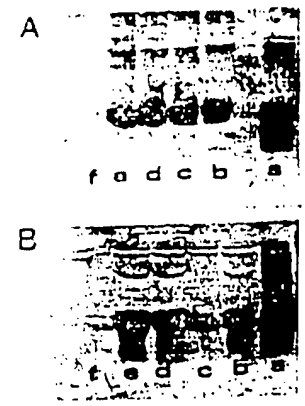


FIG. 2. **Western blot of crude cell lysates from freeze-thawed lysed *E. coli* BL21(DE3) expressing recombinant native and mutant αB-crystallin.** Bovine α-crystallin was used as a control marker. *A*: *a*, bovine α-crystallin; *b*, native αB-crystallin; *c*, F27R; *d*, F27A; *e*, K174L/K175L; *f*, negative control pET 3d. *B*: *a*, bovine α-crystallin; *b*, native αB-crystallin; *c*, D2G; *d*, F24R; *e*, K174G/K175G; *f*, negative control pET 3d.

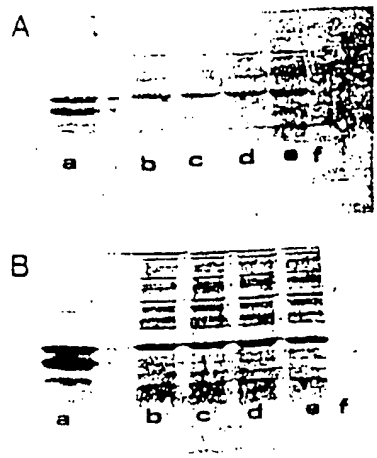


FIG. 3. **Coomassie-blue stained SDS-PAGE gels of purified native and mutant αB-crystallins.** Purification was by HPLC on a Hi-pep Sephacryl S-300 high resolution gel filtration column. See text for details. *A*: *a*, bovine αB-crystallin control marker; *b*, native αB-crystallin; *c*, F27R; *d*, F27A; *e*, K174L/K175L; *f*, control pET 3d. *B*, *a*, bovine αB-crystallin control marker; *b*, native αB-crystallin; *c*, D2G; *d*, F24R; *e*, K174G/K175G; *f*, control pET 3d.

make substitution mutations both to the C-terminal lysines (Lys^{174}, Lys^{175}) and to the N-terminal aspartate (Asp^{2}). We also made substitution mutations within the phenylalanine-rich region (Fig. 1), including F27A, which converts the αB-crystallin sequence within that region to that of the small heat-shock protein hsp27. This paper discusses the production of mutant αB-crystallins and comparison of the protein structure and chaperone-like activities of the mutant and native recombinant proteins.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmid—*E. coli* DH5α (F^-, rec^-, $meth^-$) used for propagation of plasmids, were obtained from Life Technologies, Inc. (Gibco BRL, Paisley, UK) and used as described previously (27, 28). *E. coli* BL21(DE3) were obtained from Novagen. The murine αB-crystallin cDNA cloned on plasmid pLens2–19, was kindly donated by Professor J. Piatigorsky of the National Institutes of Health. Expression plasmid pET 3d (29) was obtained from Novagen. Cloning vector pBlueScriptSK was obtained from Stratagene.

Enzymes and Media—The restriction *NcoI*, *EcoRV*, and *BamHI* endonucleases, Klenow polymerase, DNA kinase, and T4 DNA ligase were purchased from Life Technologies, Inc., *Taq* polymerase was from Perkin-Elmer Cetus (Perkin-Elmer Corp., Warrington, UK). Chemicals, including hsp27, were from Sigma and of molecular biology grade as

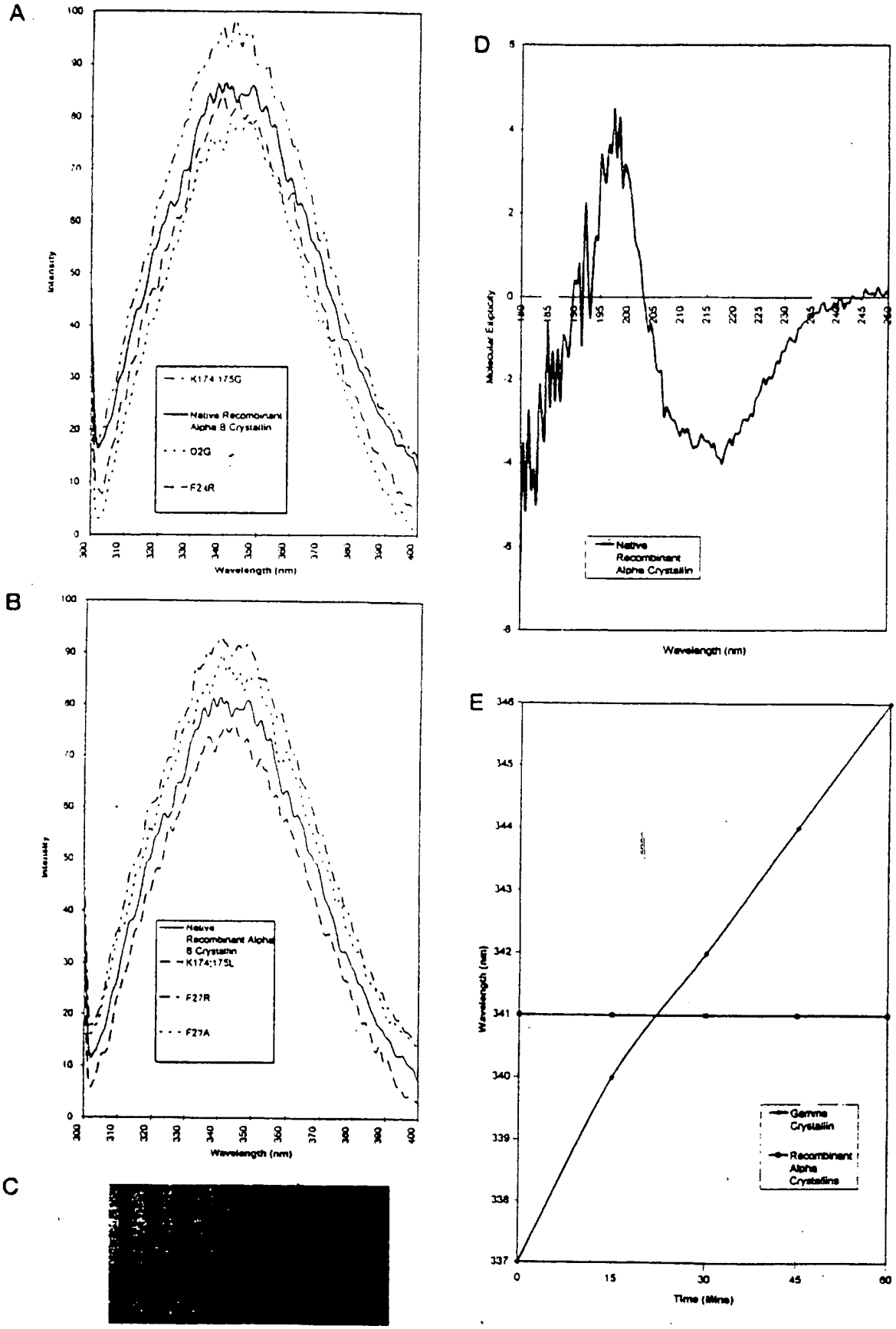

A
Intensity
Wavelength (nm)
K174 175G
Native Recombinant Alpha B Crystallin
G2G
F24R
B
Intensity
Wavelength (nm)
Native Recombinant Alpha B Crystallin
K174;175L
F27R
F27A
C
D
Molecular Ellipticity
Wavelength (nm)
Native Recombinant Alpha Crystallin
E
Wavelength (nm)
Time (Mins)
Gamma Crystallin
Recombinant Alpha Crystallins

FIG. 4

appropriate, unless otherwise stated. Cells were propagated in Luria media, and recombinant bacteria were selected using ampicillin (30).

Subcloning of Murine αB-Crystallin cDNA and Preparation of Plasmids – PCR[1] amplification (31) using primers containing *Nco*I sites was used to prepare native and mutant amplicons of murine αB-crystallin. After PCR amplification and subsequent purification, the αB-crystallin amplicons were blunt-ended with Klenow fragment DNA polymerase, phosphorylated with DNA kinase, and then cloned into *Eco*RV-cut pBlueScript SK. The pBS-SK was then cut with *Nco*I liberating αB-crystallin cDNAs flanked by *Nco*I sticky ends. These were then inserted into the *Nco*I site of pET 3d. Recombinant plasmids were identified and orientated by *Bam*HI digests. Plasmid DNA was propagated and purified by the standard methods (30).

Site-directed Mutagenesis – Site-directed mutagenesis was carried out using PCR. C-terminal substitutions were made by incorporating mismatches in the C-terminal PCR primers. Substitutions in the RLFDQFF region were made using overlap extension PCR mutagenesis (32). PCR may introduce occasional random mutations; therefore, DNA sequences of native and mutant amplicons were verified using the standard Sequenase dideoxynucleotide chain termination method (33, 34).

Expression, Purification, and Quantitation of Native and Mutant Recombinant αB-Crystallins – *E. coli* BL21(DE3) cells were transformed by the standard *E. coli* transformation procedure (30). Transformants were grown at 37 °C in Luria broth to A_{600} = 0.6, and αB-crystallin expression was then induced by addition of isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 1 mM, then the culture was incubated at 37 °C for 12 h.

Cells from 500-ml cultures were collected by centrifugation at 3,000 × *g* for 5 min at 4 °C and resuspended in 20 ml of lysis buffer (100 mM Tris-HCl, 0.05% aprotinin, 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride, 10 mM dithiothreitol, pH 7.5). The suspensions of cells were disrupted using 3 passages through a French pressure cell at 12,000 p.s.i.

The bacterial lysates were then centrifuged at 12,000 × *g* for 5 min at 4 °C, and the supernatant was assayed for the presence of αB-crystallin by SDS-PAGE and Western blotting using rabbit anti-murine α crystallin antibodies.

Nucleotides which would interfere with spectrophotometric estimation of protein concentration were precipitated from the soluble fraction by addition of polyethylenimine and dithiothreitol to final concentrations of 0.12% and 10 mM, respectively. Incubation at room temperature for 10 min was followed by centrifugation at 15000 × *g* for 10 min. The supernatant containing the proteins was then removed and the recombinant αB-crystallin was purified by HPLC gel filtration in 100 mM sodium phosphate buffer (pH 7.4) on a Pharmacia Hi-Prep Sephacryl S-300 high resolution column. Fractions containing the αB-crystallin were identified by immuno dot-blotting. The elution volume of the αB-crystallin peaks were used to estimate the size of the recombinant protein aggregates. Purity of the αB-crystallin in the positive fractions was assessed by SDS-PAGE. The purified protein concentration was then estimated by A_{280} determination using a Beckman DU-70 spectrophotometer.

N-terminal Sequencing of Recombinant αB-Crystallin – Mammalian α-crystallins have blocked N termini and cannot be sequenced; however, proteins expressed in *E. coli* frequently have unblocked N termini. Samples of recombinant αB-crystallins were therefore sent to Dr. A. Willis at the MRC Immunochemistry Unit, Oxford, for N-terminal sequencing, as described previously (27).

Aggregate M_r, Tryptophan Fluorescence, and Circular Dichroism – Mutations could affect the ability of the αB-crystallin monomers to form aggregates. Nondenaturing PAGE was used to compare the size of the αB-crystallin aggregates formed by each of the recombinant proteins.

The three-dimensional structures of native and mutant purified recombinant αB-crystallins were investigated by both circular dichroism and tryptophan fluorescence using HPLC buffer as a blank and control samples of bovine α-crystallin.

Far-UV circular dichroism spectra of each recombinant αB-crystallin were determined using an ISA JOBIN YVON CD6 Dichrograph with a 10-μm cell. Five repeat spectra were obtained for each sample and averaged out to minimize noise in the final spectrum. CD spectra were analyzed using Contin software (35) to estimate the secondary structure content of each mutant. Tryptophan fluorescence spectra (exciting at 295 nm) were determined using a Perkin Elmer L550 spectrofluorimeter.

Assays of in Vitro Chaperone-like Activity – The chaperone-like activity of the purified recombinant αB-crystallins was assayed by both the heat aggregation method (19) using γ-crystallin as substrate and the reduced insulin B chain method (36) at varying concentrations of the αB-crystallins. For the reduced insulin assay, we also included the small heat shock protein hsp27. (All spectrophotometry was carried out using a Beckman DU-70 spectrophotometer with a water-heated cuvette holder.)

Measurement of Death Rate in E. coli – Potential *in vivo* heat shock protein activity of the native and mutant recombinant αB-crystallins was determined by comparing the thermal death curves at 50 °C of stationary phase *E. coli* BL21(DE3) cells expressing the different recombinant αB-crystallins with that of *E. coli* BL21(DE3) containing only pET 3d without a cDNA insert. Bacteria were grown and induced with isopropyl-1-thio-β-D-galactopyranoside as for protein production (see above), and then 50-ml cultures were placed in 50 °C water baths for 12 h. Samples were removed at 1–2-h intervals and the numbers of surviving colony-forming units/ml were determined by the standard spread plate method on Luria/ampicillin plates.

RESULTS

Expression and Purification of Native and Mutant αB-Crystallin in E. coli Cells – Western blots of soluble fractions from *E. coli* BL21(DE3) transformed with pET 3d/αB-crystallin amplicon recombinants identified an expressed protein (absent in nonexpressing control cells) which co-migrated with control α-crystallin and cross-reacted with anti-α-crystallin antibodies (Fig. 2, *A* and *B*). N-terminal sequencing of this expressed protein demonstrated that the N terminus was not blocked and identified the first 10 residues as identical to the known sequence of murine αB-crystallin: MDIAIHHPWI. Gel filtration chromatography succeeded in purifying the expressed αB-crystallin to a purity of 85–90% (Fig. 3, *A* and *B*) with a yield of 3–4 mg/ml.

Site-directed Mutagenesis – Site-directed mutagenesis of αB-crystallin was used to produce the six mutants shown in Fig. 1. DNA sequencing was used to verify the authenticity of native and mutant recombinant αB-crystallin amplicons. All of these mutants were expressed in *E. coli* BL21(DE3) as soluble proteins with monomer size and antigenic reactions identical to that of the native and native recombinant proteins (Fig. 2).

Native Aggregate Size, CD, and Tryptophan Fluorescence of Recombinant αB-Crystallins – The native aggregate size of α-crystallins is very high (approximately 800 kDa) and difficult to estimate by PAGE or HPLC. However, comparison of the immuno dot-blot-positive fractions from the Sephacryl S-300 column demonstrated that control bovine α-crystallin, native recombinant αB-crystallin, and all 6 mutants eluted in the 94–97-ml fractions with no lower molecular weight α-crystallin species detected in later fractions. This was confirmed by the nondenaturing PAGE (Fig. 4*C*) which demonstrated no observable low molecular weight species in the purified αB-crystallin samples. In all samples (native and mutant), the protein was

[1] The abbreviations used are: PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; HPLC, high performance liquid chromatography.

Fig. 4. **Structural characterization of recombinant αB-crystallins.** *A* and *B*, tryptophan fluorescence spectra of purified native and mutant recombinant αB-crystallins. Expressed α-crystallins were excited at 295 nm. All gave identical emission maxima at 340–341 nm. Bovine α-crystallin was used as a control. Intensities were all identical and have been vertically displaced for clarity. *C*, Coomassie blue-stained nondenaturing PAGE gel of native and mutant αB-crystallins. The *break* in each lane is the junction between stacking and separating gels. *a*, bovine α-crystallin control marker; *b*, native αB-crystallin; *c*, F27R; *d*, F27A; *e*, K174*L*/K175L; *f*, F24R; *g*, D2G; *h*, K174G/K175G; *i*, bovine α crystallin control marker. *D*, circular dichroism of recombinant native αB-crystallin. All mutant recombinant αB-crystallins showed identical CD spectra. *E*, changes in tryptophan fluorescence emission spectra of purified recombinant αB-crystallins and γ-crystallin at 66 °C.

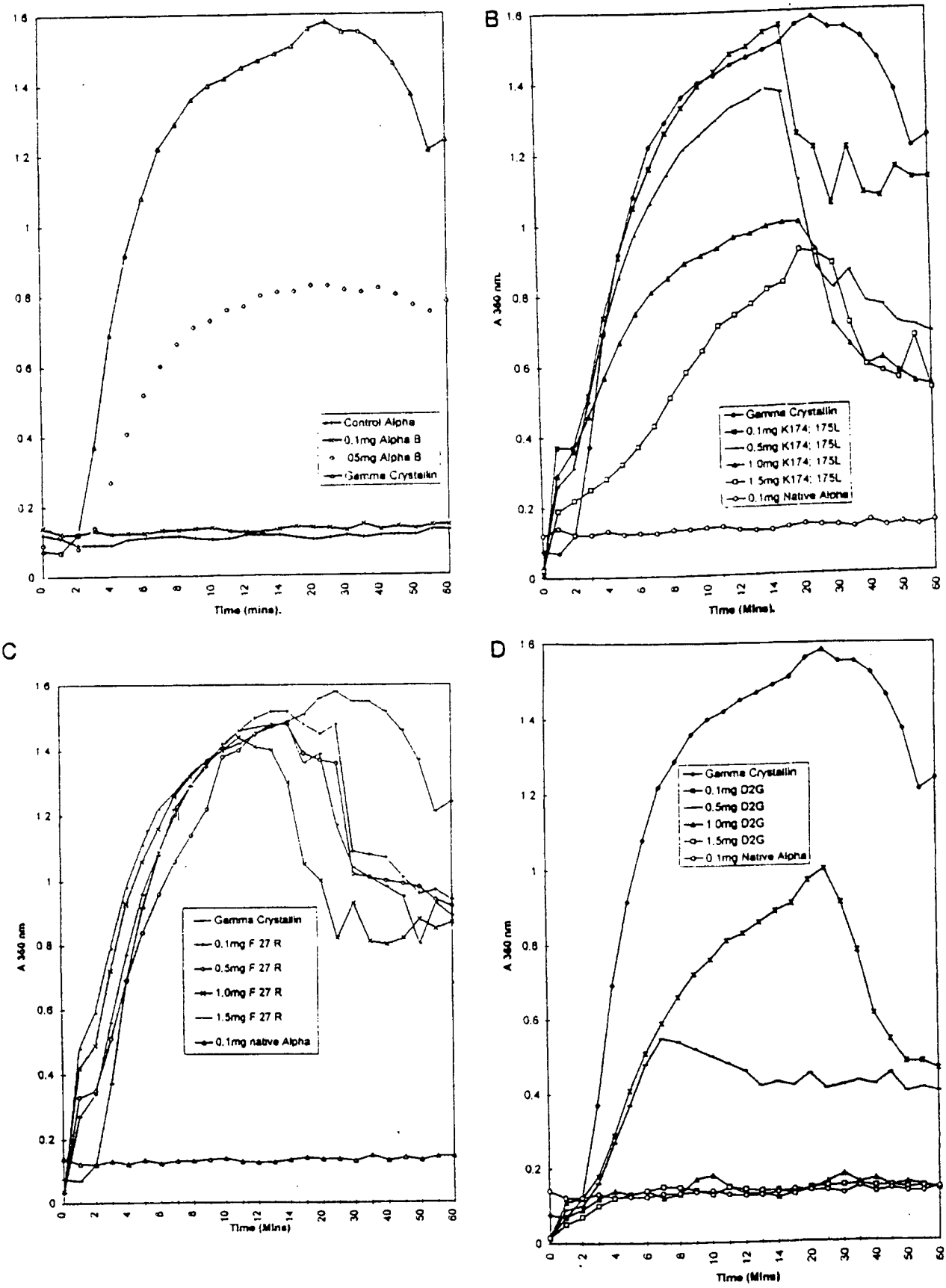


FIG. 5. **Inhibition of γ-crystallin aggregation assay in the presence and absence of variable amounts of recombinant native and mutant αB-crystallin at 66 °C.** *A*, native recombinant αB-crystallins; *B*, mutant K174L/K175L αB-crystallins; *C*, mutant F27R αB-crystallins; *D*, mutant D2G αB-crystallins.

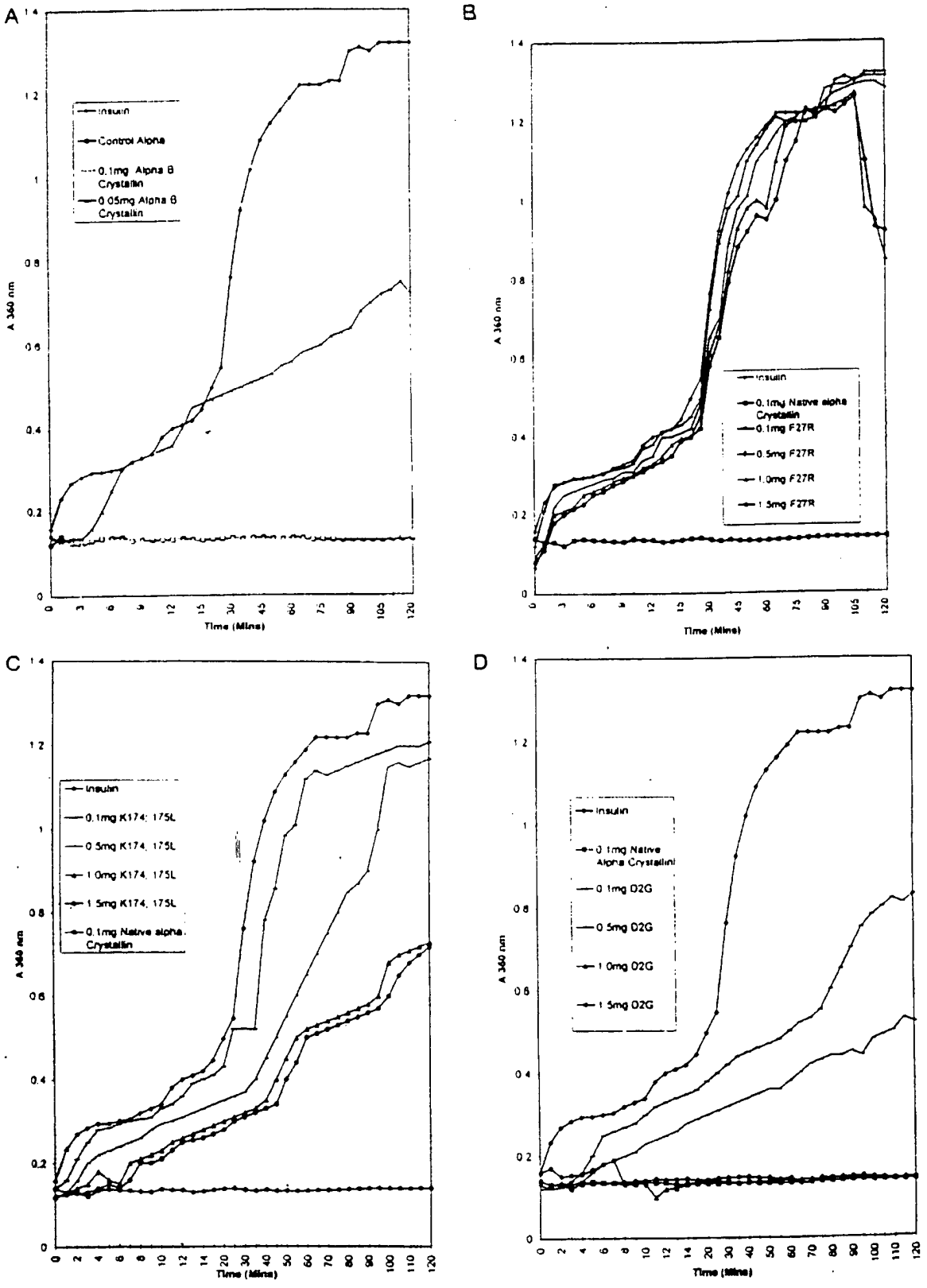


FIG. 6. **Aggregation of the insulin B chain by reduction with dithiothreitol in the presence and the absence of variable amounts of recombinant native and mutant αB-crystallin.** *A*, native recombinant αB-crystallins; *B*, mutant F27R; *C*, mutant K174L/K175L; *D*, mutant D2G αB-crystallins.

found as a high M_r species which sat on the top of the 5% separating gel. Thus, it appears that all of the mutations studied did not alter the aggregation behavior of the αB-crystallin in such a way as to produce monomers, tetramers, or other low molecular weight species.

Circular dichroism studies revealed that the recombinant native and mutant αB-crystallins contained 42% β structure ± 3–5% and 12% α helix ± 1–2%, which represent no significant differences in secondary structure (Fig. 4*D*). Tryptophan fluorescence emission spectra from native and mutant recombinant α-crystallins all gave emission maxima at 341 nm (Fig. 4, *A* and *B*). This was identical with emission maxima of the positive control of bovine α-crystallin. The emission maxima were consistent with previously documented spectra (37). Thus, both circular dichroism and tryptophan fluorescence studies show that the mutations had not produced any gross significant structural alteration.

Tryptophan fluorescence emission maxima were also observed during prolonged incubation at 66 °C to evaluate any possible structural changes that could occur at elevated temperatures similar to those that exist in the γ-crystallin aggregation assay. A control sample of bovine γ-crystallin was observed to undergo significant unfolding as demonstrated by a shift in the emission maximum from 337 to 346 nm. There were no changes in the emission maxima after a 1-h incubation at 66 °C indicating no heat-induced structural changes in any of the α-crystallins (Fig. 4*E*).

In Vitro Assessment of Chaperone-like Activity of Native and Mutant αB-Crystallins—Native recombinant αB-crystallin was shown to possess chaperone-like activity. 0.1 mg of native αB-crystallin successfully inhibited the thermal aggregation of γ-crystallin. 0.05 mg of αB-crystallin reduced the chaperone-like activity by 50% (Fig. 5*A*). Mutant αB-crystallins showed no significant chaperone-like activity (Fig. 5, *B* and *C*), apart from the D2G mutant, which demonstrated a reduced efficiency in chaperone-like activity (Fig. 5*D*).

A similar pattern of chaperone-like activity was obtained for each recombinant protein in the room temperature reduced insulin B chain aggregation assay (Fig. 6).

In contrast, 0.1 mg of hsp27 failed to exhibit any chaperone-like activity in the reduced insulin assay, in a fashion similar to F27R (Fig. 5*C*).

In Vivo Assessment of Chaperone-like Activity of Native and Mutant αB-Crystallins—Native αB-crystallin was shown to confer thermotolerance on and preserve longevity of *E. coli* BL21(DE3), when compared to pET 3d control cells. Mutant αB-crystallins failed to confer thermotolerance to *E. coli* BL21(DE3) apart from the D2G mutant which conferred a similar but reduced thermotolerance to that conferred by native αB-crystallin (Fig. 7 and Table I).

DISCUSSION

Recombinant αB-crystallin lacks the blocked N terminus found in the native lens protein but appears to be similar in all other respects (tryptophan fluorescence, circular dichroism, oligomer size, and chaperone-like activity) to the lens protein. This implies that the presence of a free positive charge on the N terminus does not influence gross α-crystallin structure or inhibit protein binding.

We have produced mutants of this protein in three specific areas: substituting neutral or charged residues for hydrophobic residues in the conserved phenylalanine-rich region, substituting neutral or hydrophobic residues for the C-terminal lysine residues, and substituting glycine for the N-terminal aspartate Asp^2.

Biophysical characterization of the expressed proteins demonstrates no global structural changes in any of the mutants studied. Oligomer size, far UV CD spectra, and tryptophan emission maxima were not significantly altered by any of the mutations. Das and Surewicz (38) have observed changes in α-crystallin structure at elevated temperatures; however, our mutations do not appear to have influenced the sensitivity of the recombinant α-crystallins to temperature. All of the recombinant α-crystallins demonstrated no shift in tryptophan emission maximum after a 1-h incubation at 66 °C, suggesting that the heat stabilities of the recombinant mutant proteins were also unaffected by the mutations. Furthermore, the results of the room temperature assay demonstrates no observable differences in mutant behavior from the 66 °C assay. Thus, observed changes in chaperone-like behavior of these mutants is likely to be a direct result of substituting key residues in the peptide binding site(s).

It is of interest that in the three functional assays we used, two *in vitro* (heat aggregation of γ-crystallin and aggregation of insulin at room temperature) and one *in vivo* (*E. coli* heat tolerance), the effects of the mutations were identical, and followed the pattern: native > D2G > K174G/K175G = K175A/K175L ≫ F24R = F27A = F27R. Substitution of the either of the phenylalanines Phe^{24} or Phe^{27} appears to completely abolish chaperone-like activity, as measured by the γ-crystallin, insulin, and *in vivo* assays. This suggests that these highly conserved hydrophobic residues play a vital role in the chaperone-like activity of αB-crystallin.

Unexpectedly, F27A, which converted the αB-crystallin sequence in the phenylalanine-rich region to that of the functional small heat shock protein hsp27, produced a protein which failed to exhibit chaperone-like activity. When hsp27 was itself used as a control in the reduced insulin assay, no chaperone-like activity was demonstrated. This suggests that the conditions of even the least aggressive *in vitro* assay were too extreme for demonstration of heat shock protein functionality. The observation that native αB-crystallin can exhibit chaperone-like activity under these conditions suggests that it may be more efficient than hsp27 in binding unfolded protein. F27A appears to abolish this increased efficiency, suggesting that phenylalanine 27 plays a key role in binding unfolded protein.

Smulders *et al.* (39) suggested that the hydrophobic residues Phe^{74}, Leu^{77}, and Val^{72} were not involved in αA-crystallin function; however, they did not investigate the RLFDQFF region. Removal of the N terminus from αA-crystallin removes a very hydrophobic region αA-(32–37) and may be responsible for the chaperone-like activity (40). This is in accord with our experiments on the interactions of α-crystallin with chymosin; binding only occurs with unfolded chymosin or with prochymosin (which contains a hydrophobic N-terminal region), not with correctly folded chymosin (41).

Mutations to the C-terminal lysines greatly reduced the chaperone-like activity of the αB-crystallin such that it failed to confer thermotolerance on *E. coli* and *in vitro* inhibition of both γ-crystallin and insulin B chain aggregation was incomplete even at very high concentrations of α-crystallin (partial protection was observed only at concentrations 15-fold greater than that required for protection by the native recombinant). This is consistent with the observations of Boyle and Takemoto (25), who suggested that the C terminus of α crystallin monomers were located in the central region where they had previously demonstrated γ-crystallin binding and were therefore likely to be involved in protein binding. It is of interest that modeling studies (42) show the C-terminal amino acids (KK) form a strong electropositive region, which is preceded by an electronegative region. The C-terminal arm could then act like a charged "fishhook" to interact with unfolded proteins via char-

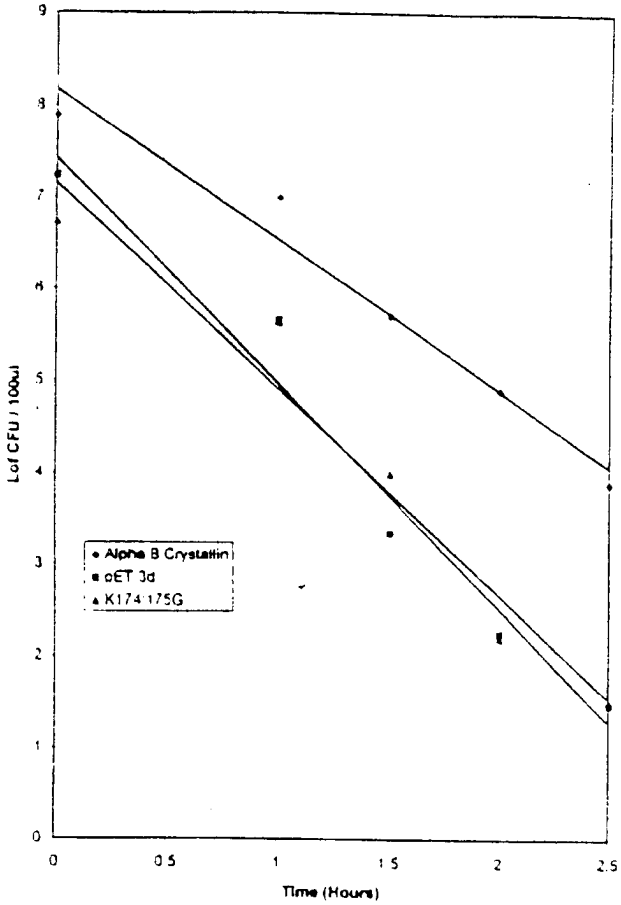


FIG. 7. **Thermal death rates of *E. coli* BL21(DE3) at 50 °C, represented as *CFU* (colony-forming units)/100 μl.** Native expressed αB-crystallins reduced the thermal death rate of *E. coli* BL21(DE3) at 50 °C compared to that observed in pET 3d control cells and in *E. coli* expressing mutant K174G/K175G under the identical conditions. Similar results were obtained with all other recombinant mutant αB-crystallins.

TABLE I
Thermal death rates of E. coli BL21(DE3) at 50 °C, represented as CFU (colony-forming units)/100 μl, expressing native and mutant αB-crystallins

	log CFU/100 μl/h
F27R	−2.7
K174L/K175L	−2.6
Negative control (no α-crystallin)	−2.4
K174G/K175G	−2.25
F27A	−2.2
F24R	−2.19
D2G	−1.9
Native αB-crystallin	−1.64

ge-charge interactions and then link the substrate proteins further via hydrophobic interactions with the exposed phenylalanine-rich domain near the N terminus. Thus, unless there were significant exposed hydrophobic residues on the surface of the substrate protein, there would not be a stable interaction with αB-crystallin and its substrate protein. Thus, even if the lysine residues had been deleted or glycated, there would be sufficient charge-charge interactions with the remainder of the C-terminal arm to ensure efficient chaperone-like activity. This is in agreement with experiments where proteolytic removal of the C-terminal ($Thr^{171} \rightarrow Lys^{175}$) region did not significantly affect αB-crystallin chaperone-like function (23). In addition, it may be that the mutations we have made resulted in a largely intact but less mobile C terminus. The absence of a strong hydrophilic positive charge at the end of the highly flexible C-terminal extension may result in the C terminus folding back on itself, losing its flexibility, and sterically hindering protein binding.

Substitution of the N-terminal aspartate (D2G) resulted in a greatly reduced efficiency of chaperone-like activity. Some thermotolerance was conferred on *E. coli* by expression of the D2G mutant, but this was significantly less than that conferred by the native recombinant protein. Similarly, in the *in vitro* assays, complete aggregation inhibition was demonstrated by the D2G mutant but only at concentrations 10-fold greater than the native recombinant protein. This suggests that while the Asp^2 residue is not vital for binding of unfolded proteins, it does play some role in the chaperone-like activity. Boyle and Takemoto (25) have suggested that the N termini of the α monomers may also be located in the binding site, and it is possible the the Asp^2 is involved in a salt bridge with the C-terminal lysines. However, in that case, one might have expected similar efficiency on *in vivo* and *in vitro* chaperone-like activity for the D2G and K174G/K175G which was not the case.

We have used three different assays to demonstrate consistent changes in chaperone-like behavior produced by specific mutations in recombinant murine αB-crystallin. The results of these assays suggest that charge-charge interactions involving the C-terminal lysines and possibly Asp^2 are important in binding of unfolded protein but also that hydrophobic interaction involving the conserved phenylalanine-rich region plays a vital role in the chaperone-like activity. This region is conserved in a number of heat shock proteins, suggesting that similar hydrophobic interactions may be involved in the activities of all small heat shock proteins, although subtle sequence modifications, such as F27A in human hsp27 (26) appear to modify the relative efficiency of binding to unfolded proteins. It is apparent that *in vitro* assays for functionality of hsps and αB-crystallin are as yet rather insensitive in recognizing cellular function(s), and investigations on novel cellular assays are under way in our laboratory.

Acknowledgments – We thank Joram Piatigorsky of the NIH for his generous gift of the murine αB-crystallin cDNA, Professor Joe Horwitz for advice on removal of nucleotides and other useful conversations, Tony Willis at the MRC Immunochemistry Unit in Oxford for the N-terminal sequencing, Ian Summner and Drummond Smith of the IFR in Reading for help with the CD and tryptophan fluorescence, and Audrey Coop for technical support.

REFERENCES

1. Wistow G., and Piatigorsky J. (1988) *Annu. Rev. Biochem.* **57**, 479–504
2. Lubsen N. H., Aarts H. J. M., and Schoenmakers J. G. G. (1988) *Prog. Biophys. Mol. Biol.* **51**, 47–76
3. Delaye M., and Tardieu A. (1983) *Nature* **302**, 415–417
4. Harding J. J., and Crabbe M. J. C. (1984) in *The Eye* (Davson, H., ed) Vol. 1B, pp. 207–492, Academic Press, New York
5. Harding J. J., and Crabbe M. J. C. (1991) in *Post-translational Modifications of Proteins* (Harding, J. J., and Crabbe, M. J. C., eds) CRC Press, London
6. Wolff S. P., and Crabbe M. J. C. (1985) *Biochem J.* **226**, 625–630
7. Iwaki, T., Kume-Iwaki, A., Liem, R. K. H., and Goldman, J. E. (1989) *Cell* **57**, 71–78
8. Klemenz, R., Frohli, E., Steiger, R. H., Schafer, R., and Ayoama, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3652–3656
9. DasGupta, S., Hohman, T. C., and Carper, D. (1992) *Exp. Eye Res.* **54**, 461–470
10. Klemenz, R., Frohli, E., Aoyama, A., Hoffman, S., Simpson, R. J., Montz, R. L., and Schafer, R. (1991) *Mol. Cell Biol.* **11**, 803–812
11. Iwaki, T., Iwaki, A., Tateishi, J., and Goldman, J. E. (1994) *J. Cell Biol.* **125**, 1385–1393
12. Renkawek, K., De Jong, W. W., Merck, K. B., Frenken, C. W. G. M., van Workum, F. P. A., and Bosman, G. J. C. G. M. (1992) *Acta Neuropathol.* **83**, 323–327
13. van Noort, J. M., van Sechel, A. C., Bajramovic, J. J., El Ouagmiri, M., Polman, C. H., Lassmann, H., and Ravid, R. (1995) *Nature* **375**, 798–801
14. Goode, D., and Crabbe, M. J. C. (1995) *Comput. Chem.* **19**, 65–74
15. Harding, J. (1991) *Cataract-Biochemistry and Epidemiology*, Chapman and Hall, London

16. Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georopoulous, C. P., Hendrix, J. W., and Ellis, R. J. (1988) *Nature* **333**, 330–339
17. Goloubinoff P., Gatenby A. A., and Lorimer G. H. (1989) *Nature* **337**, 44–47
18. Zahn, R., Perrett, S., Stenberg, G., and Fersht, A. R. (1996) *Science* **271**, 642–645
19. Horwitz, J. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10449–10453
20. Takemoto, L., Emmons, T., and Horwitz, J. (1993) *Biochem. J.* **294**, 435–438
21. Crabbe, M. J. C., Plater, M. L., and Goode, D. (1995) *Proceedings of the International Forum on Advanced Techniques in Lens Cataract Research*, 28, Kanazawa, Japan
22. Rao, P. V., Huang, Q., Horwitz, J., and Zigler, J. S., Jr. (1996) *Biochim. Biophys Acta* **1245**, 439–447
23. Carver, J. A., Guerreiro, N., Nicholls, K. A., and Truscott, R. J. W. (1995) *Biochim. Biophys. Acta* **1252**, 251–260
24. Bettelheim, F. A., Reid, M. B., McPhie, P., and Garland, D. (1992) *Biochem. Biophys. Res. Commun.* **187**, 39–44
25. Boyle D., and Takemoto L. (1994) *Exp. Eye Res.* **58**, 9–16
26. Crabbe, M. J. C., and Goode, D. (1994) *Biochem J.* **297**, 653–654
27. Goode, D., and Crabbe, M. J. C. (1994) *Arch. Biochem. Biophys.* **315**, 104–110
28. Dilsiz, N., and Crabbe, M. J. C. (1995) *Biochem. J.* **305**, 753–759
29. Barbosa P., Cialkowski M., and O'Brien W. E. (1991) *J. Biol. Chem.* **266**, 5286–5290
30. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
31. Saiki, R. K., Gelfland, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science* **239**, 487–491
32. Horton, R. M., and Pease, L. R. (1991) in *Directed Mutagenesis: A Practical Approach* (McPherson, M. J., ed) pp. 217–248, IRL Press, Oxford
33. Sanger, F., Niklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
34. Dilsiz, N., and Crabbe, M. J. C. (1994) *Anal. Biochem.* **222**, 510–511
35. Provencher, S. W. (1982) *Comp. Phys. Comm.* **27**, 229–242
36. Farahbakhsh, Z. Y., Huang, Q.-L., Ding, L.-L., Altenbach, C., Steinhoff, H.-J., Horwitz, J., and Hubbell W. L. (1995) *Biochemistry* **34**, 509–516
37. Palmineo, D. V., Groth-Vaselli, B., Fansworth, P. N., and Reddy, M. C. (1995) *Biochim. Biophys. Acta* **1246**, 91–97
38. Das, K. P., and Surewicz, W. K. (1995) *FEBS Lett.* **369**, 321–325
39. Smulders, H. P. H., Merck, K. B., Aandekerk, J., Horwitz, J., Takemoto, L., Slingsby, C., Bloemendal, H., and De Jong, W. W. (1995) *Eur. J. Biochem.* **232**, 834–838
40. Smith, J. B., Liu, Y., and Smith, D. L. (1996) *Exp. Eye Res.* **6**, 125–127
41. Chitpinityol, S., Goode, D., and Crabbe, M. J. C. (1996) in *Perspectives on Protein Engineering* (Geisow, M., ed) p. 29, Montpellier, France
42. Crabbe, M. J. C., Appleyard, J. A., and Rees Lay, C. (1994) *DeskTop Molecular Modeller*, version 3.0, Oxford University Press, Oxford, UK

The Patent Office

Application No:	GB 9722594.0	**Examiner:**	Colin Sherrington
Claims searched:	1 to 10	**Date of search:**	16 December 1997

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.O): C3H(HB4A,HB4B,HB5); G1B(BAG,BAH)

Int Cl (Ed.6): C12Q 1/68

Other: ONLINE: WPI,CLAIMS,DIALOG/BIOTECH

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	WO 95/35505 A1 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) -whole document	1, 5
A	Science 1995,270,467-470 -Mark Schena *et al.* "Quantitative Monitoring of Gene Expression with a Complementary DNA Microarray"	1, 5
A	Nucleic Acids Research 1994,22(24),5456-5465 -Zhen Guo *et al.* "Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports"	1, 5
A/X	Analytical Biochemistry 1993,209,63-69 -Shintaro Kawai *et al.* "A Simple Method of Detecting Amplified DNA with Immobilized Probes on Microtiter Wells"	1 / 5 to 10

X Document indicating lack of novelty or inventive step
Y Document indicating lack of inventive step if combined with one or more other documents of same category.
& Member of the same patent family
A Document indicating technological background and/or state of the art.
P Document published on or after the declared priority date but before the filing date of this invention.
E Patent document published on or after, but with priority date earlier

9. An array as claimed in any one of claims 5-7 wherein the probe DNAs are sense strand DNA.

10. The use of an array as claimed in any one of claims 5-9 to provide a quantitative estimate of the abundance of individual mRNAs or their corresponding first strand cDNAs within a complex mixture of such derived from a biological sample comprising a single cell type or a mixed population of cell types.

CLAIMS:

1. A method for preparing an array of single-stranded DNA immobilised on a solid support, which method comprises (i) providing samples of double-stranded DNA chemically modified on the sense or antisense strand for attachment to the solid support, and (ii) linking the DNAs to the solid support and, before or after step (ii), removing the non-modified strand whereby an array of single-stranded DNA is provided on the solid support.

2. A method as claimed in claim 1 wherein the single-stranded DNA comprises DNA molecules containing more than 75 nucleotides.

3. A method as claimed in claim 1 or claim 2 wherein the strand that is not to be bound to the solid support is chemically modified to assist strand separation or its selective degradation.

4. A method as claimed in any one of the previous claims wherein the double-stranded DNAs are the amplification products of chemically modified polymerase chain reaction (PCR) primers.

5. An array of single stranded probe DNAs, each probe comprising at least 75 nucleotides and being chemically immobilised on a solid support.

6. An array as claimed in claim 5 wherein each probe comprises at least 200 nucleotides.

7. An array as claimed in claim 5 or claim 6 wherein the probe DNAs are of unknown sequence.

8. An array as claimed in any one of claims 5-7 wherein the probe DNAs are antisense strand DNA.

The slide is incubated in a humdified atmosphere at 37_C for 2h, washed with 1% NH_4OH and water and air dried at room temperature.

Glass slides containing arrays of paired elements comprising sense and antisense probes are hybridised with first-strand cDNA prepared by reverse transcription of polyA mRNA isolated from HepG2 cells and labelled with the fluorescent nucleotide analogue dCTP-Cy5 (Amersham International, Chalfont, UK) essentially as described by Schena et al. [Science, 1995, 270, 467-470]. The labelled cDNA (5 micrograms in 7.5 microlitres) is denatured at 95°C. 2.5 microlitres of concentrated hybridisation solution (5 x SSC, 0.1% SDS) is added and the mixture transferred to the glass microscope slide over the array under a cover slip. Hybridisations are carried out in a humidified atmosphere for 12h at 65°C, and the slides washed twice in 0.1 x SSC at 60°C. Fluoresence detection and image reconstruction is carried out as described by Guo et al. [Guo, Z. et al., Nucleic Acids Research, 1994, 22, 5456-5465].

Example 1

Array elements were selected from a set of clones isolated from a human liver cDNA library containing cDNA inserts cloned unidirectionally into a pBluescript vector (Stratagene) between the EcoRI and XhoI sites, such that the 3' end of the insert DNA, including the polyA tail, is located immediately adjacent to the XhoI site. The library was maintained in E. coli strain SOLR™ (Stratagene). The average insert size was 1.5kb. Randomly selected clones were transferred to a 96-deep well microtitre plate and grown in L-broth supplemented with 100 micrograms/ml ampicillin.

To generate sense probes the DNA inserts were amplified using a pair of 24nt primers corresponding to the vector sequences immediately flanking the two restriction sites. The sense primer complementary to pBluescript sequences 5' to the EcoRI site was synthesised with a 6-aminohexyl-phosphodiester at its 5' end. The antisense primer complementary to pBluescript sequences 3' to the XhoI site was biotinylated at its 5' end according to standard procedures [Agrawal, S. et al., Nucleic Acids Research, 1986, 14, 6227-6245]. The PCR reactions with the modified primers were performed directly on small volumes (typically <1 microlitre of overnight culture) of the bacterial cultures in a 96-well thermocycler in final reaction volume of 70 microlitres. Each of the PCR products was purified using a QIAquick™ PCR purification kit (Qiagen Inc., Chatsworth, CA).

Strand separation and removal of the antisense strands was carried out as previously described [Guo, Z. et al., Nucleic Acids Research, 1994, 22, 5456-5465]. The remaining sense strand 5'amino-modified probes were dried down in vacuo and redissolved in 20 microlitres of 100mM sodium carbonate/bicarbonate buffer (pH9.0)

To generate antisense strand probes, the PCR reactions were carried out as described above using an antisense primer with a 5' 6-aminohexyl-phosphodiester group and a biotinylated sense primer. The resulting products were purified and strand-separated as described above.

Pre-cleaned glass microscope slides are treated with 1% 3-aminotrimethoxysilane solution (Aldrich Chemical, Milwaukee, WI), washed, dried and activated with 1,4-phenylene di-isothiocyanate as described by Guo et al. [Guo, Z. et al., Nucleic Acids Research, 1994, 22, 5456-5465]. 2 microlitre samples of either the sense strand or antisense strand DNAs (typically 0.25 - 0.5 micrograms of DNA) are spotted manually onto the microscope slide.

measured using an array of elements comprising antisense single-strand probes, and non-specific hybridisation is measured using an array of the corresponding sense single-strand probes. These two sets of probes may be immobilised on the same or separate solid surfaces as described above. In the case of labelled cDNAs, specific hybridisation is measured using an array of elements comprising sense single-strand probes, and non-specific hybridisation is measured using an array of the corresponding antisense single-strand probes. These two sets of probes may be immobilised on the same or separate solid surfaces as described above.

Hybridisation conditions at the solid support will depend on the nature of the support and the arrayed DNA, but may be defined and optimised using a number of methodologies available to one ordinarily skilled in the art [see e.g., Sambrook, J. et al., "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989]. Preferably, hybridisation takes place under stringent conditions, i.e. those which reveal nucleic acid identities of greater than 95%. However, if desired, other less stringent hybridisation conditions may be selected. Hybridisation of a particular nucleic acid species is detected by measuring the strength of the signal from the labelled target nucleic acid that remains bound its cognate element in the array after washing the array at the particular stringency chosen for the application.

The absolute abundance of a particular single-strand nucleic acid species (be it mRNA or first strand cDNA) in a plurality of nucleic acids may be determined by subtracting the signal at the element in the array corresponding to the non-specific hybridisation from the signal at the element in the array affording the specific hybridisation signal for that particular nucleic acid species. To determine whether the expression of a particular mRNA is altered in some condition, for example a diseased state compared to the normal state, identical arrays are hybridised to labelled samples of target nucleic acids isolated from the diseased and normal biological samples. Differences in the measured abundance can be used to indicate which genes may be involved in the cause, maintenance or progression of the chosen diseased state. The same approach can be used to follow the effects of drug treatment or other investigation of or manipulation of a set of cells or an organism on the expression levels of the genes within the biological sample.

electrostatic interactions, such as the binding of probe DNA to poly-L-lysine coated slides, where portions of the DNA probe are complexed with the poly-L-lysine and therefore not available for hybridisation to target DNA. The quantity of single-strand DNA that is arrayed at each element in the array and is free to hybridise to target nucleic acid will vary according to the nature of the solid surface and the chemistry used to link the probes to the solid surface. However it will be present in sufficient quantity to ensure that it is always in excess relative to the concentration of its corresponding labelled target nucleic acid in the sample to be analysed. In this way, the intensity of the resultant hybridisation signal will be proportional to the amount of target nucleic acid present in the biological sample.

In a further aspect of this invention we provide sense strand arrays which comprise a plurality of DNA elements comprising sense strands immobilised on a single solid surface, where the strands in each element are derived from a different polynucleotide DNA fragments and are prepared according to the method described above. In the same way, we provide antisense strand arrays which comprise a plurality of DNA elements comprising antisense strands immobilised on a single solid surface, where the strands in each element are derived from a different polynucleotide DNA fragment and are prepared according to the method described above

In a preferred embodiment of this invention, mixed arrays can be constructed containing pairs of elements comprising either the sense or antisense strand of a given DNA. The pairs of elements do not necessarily have to be arrayed side-by-side within the array. The precise disposition of the two types of element, either within the same array or on different arrays will depend on the precise application for which they are intended.

iv) Hybridisation

Selected single-strand arrays generated as described above may be hybridised to a sample containing a plurality of single-strand target nucleic acids, either mRNAs or preferably, first strand cDNAs that have been isolated from a chosen biological sample and labelled by any of the techniques known to one ordinarily skilled in the art, such as radiolabelling, fluorescent labelling or chemiluminescent labelling [see e.g., Schena M. et al., Science, 1995, 270, 467-470]. In the case of labelled mRNAs, specific hybridisation is

separated by treatment with 0.1N NaOH for 10 minutes. The beads containing the bound unwanted strand are removed by centrifugation and the supernatant containing the desired non-biotinylated strand is decanted and neutralised to pH7.0 with HCl. This strand is then arrayed and bound to the solid support through the functionality it carries at its 5' end; for example a 5' aminohexyl-phosphodiester group which will couple to glass activated with 1,4-phenylene di-isothiocyanate (DITC).

In a second embodiment, the double-stranded PCR product is arrayed first and the chosen strand coupled to the solid support using the desired chemistry incorporated into the appropriate PCR primer. For example, a strand containing a 5' aminohexyl-phosphodiester group can be coupled to DITC-activated glass. In this embodiment, the unwanted strand will be unable to couple to the solid support because it has been generated using a PCR primer which lacks a 5' amino group. The arrayed double-stranded probe is then denatured, for example using a bath containing 0.1N NaOH, and the unwanted strand washed off. The solid support is then placed in a neutralising bath at pH7.0 to generate the selected strand array.

In a further embodiment, the double-stranded PCR product is arrayed first and the chosen strand linked to the solid support using the desired chemistry incorporated into the PCR primer for that strand. In this embodiment, the unwanted strand is synthesised using a PCR primer which carries an unmodified 5' phosphate group. This strand is then enzymically degraded using a 5' -3' exonuclease, for example lambda exonuclease which cannot attack 5' ends unless they carry a 5' terminal phosphate group ["Current Protocols in Molecular Biology", Ausubel, F. M. et al. (Eds.), Green/Wiley, New York, 1995, pp15.2.5].

In a further embodiment of this invention, the unwanted strand may be removed by a combination of enzymic degradation followed by alkaline denaturation, washing and neutralisation. This combination is particularly effective for probes derived from polynucleotide DNA fragments of lengths approaching 10kb.

Covalently coupling the probe DNA to the solid support at each elemental position through a 5' chemical linker has several advantages. It ensures a robust linkage of DNA to the solid surface which will be resistant to chemical degradation during storage and subsequent procedures, with consequent loss of signal. Importantly, it also provides the maximal amount of single-strand probe DNA which is free to hybridise to target DNA sequences. This is an important advantage over methods that rely on non-specific

strand is designed to facilitate subsequent strand separation, as described below. For example, where the sense strand primer contains a 5' 6-aminohexyl-phosphodiester group, the antisense strand primer could contain either a 5' phosphate or a 5' biotinylated nucleotide derivative.

In the case of an antisense strand probe the primer used to direct the DNA-polymerase dependent synthesis of the antisense strand contains a first chemical modification which will be used to couple that strand to the solid support. In this instance, the corresponding sense strand primer is either unmodified, or contains a second chemical modification different from that used in the antisense primer. In this embodiment, the modification carried on the sense strand is designed to facilitate subsequent strand separation, as described below. For example, where the antisense strand primer contains a 5' 6-aminohexyl-phosphodiester group, the sense strand primer could contain either a 5' phosphate or a 5' biotinylated nucleotide derivative.

PCR reactions are carried out on DNA obtained from individual clones to obtain the desired number of 5' modified polynucleotide DNA fragments that are to be used as probes in the selected strand arrays. These reactions may be efficiently carried out in high numbers using samples in 96- or 384-well plates and thermocyclers specifically designed to handle such plates. The PCR conditions required for each template will depend upon the precise application and can be readily optimised by anyone ordinarily skilled in the art. The products may be partially purified to remove salts, excess primers and excess nucleotides using an appropriate purification medium such as Sephacryl S-200 which will remove low molecular weight components from the PCR mix.

iii) Preparation of Single-Strand Arrays

The PCR products prepared using modified sense or antisense primers may be separated into sense and antisense strands in two ways. In the first embodiment, the two strands are separated prior to arraying onto the solid support. One desirable method to achieve this is to generate PCR products in which the unwanted strand contains one or more biotinylated nucleotides at the 5' end [Guo, Z. et al., Nucleic Acids Research, 1994, 22, 5456-5465]. The PCR product is bound to streptavidin-coated agarose beads which may be washed to remove other reagents such as salts, primers and free nucleotides, and the two strands then

complementary to the vector sequences immediately flanking the insert DNA sequence. In this way it is not necessary to know the sequence of the polynucleotide DNA fragment comprising the insert in order to practise the invention.

To prepare a selected single-strand probe, the primer used to direct DNA-polymerase dependent synthesis of the selected strand contains a first chemical modification which will be used to couple that strand to the solid support. Many methods are known in the art whereby nucleic acids can be immobilised on a variety of solid surfaces. In its preferred embodiment, the chemical modification will be incorporated into the 5' nucleotide of the primer at either the 5' phosphate, the 5' deoxyribose group or the 5' base (adenine, guanine, thymidine or cytosine) during synthesis of the oligonucleotide. However, the modification may also be made at other positions within the 5' primer sequence. The modification comprises a chemical functionality for binding to the solid surface, together with a spacer group of appropriate length to improve the accessibility of the probe to the target nucleic acid [see e.g., Maskos, U. and Southern, E. M., Nucleic Acids Research, 1992, 20, 1679-1684 for a discussion of factors influencing linker design]. In one embodiment, the chemical functionality may direct non-covalent binding to the solid surface, for example a biotin moiety which will interact with a streptavidin coating on the solid surface. In an alternative embodiment, the chemical functionality may covalently couple the selected DNA strand to the solid surface. There are a number methods for covalently attaching DNA to solid surfaces through the introduction of various chemical functional groups [see e.g. Ghosh, S. S. and Musso, G. F., Nucleic Acids Research, 1987, 15, 5353-5372; Bischoff, R. et al., Analytical Biochemistry, 1987, 164, 336-344; Guo, Z. et al., Nucleic Acids Research, 1994, 22, 5456-5465]. The precise choice of chemical functionality to be employed will depend on the nature of the solid surface onto which the DNA is to be immobilised. The spacer group may be, for example, a long-chain hydrocarbon of general formula $-(CX_2)n-$ where X may be H or F and n is generally 6-20.

In the case of a sense strand probe the primer used to direct the DNA-polymerase dependent synthesis of the sense strand contains a first chemical modification which will be used to couple that strand to the solid support. In this instance, the corresponding antisense strand primer is either unmodified, or contains a second chemical modification, different from that used in the sense primer. In this embodiment, the modification carried on the antisense

property. Where the solid support is porous, the term "solid support" refers without distinction to a range of pore sizes, depending upon the nature of the system.

As used herein, the term "surface" means any generally two-dimensional structure on a solid support to which the desired probe DNA is attached or immobilised.

As used herein, the term "target" refers to any complex mixture of nucleic acid or any individual component thereof which can be labelled such as to permit its detection by anyone ordinarily skilled in the art.

As used herein, the term "vector" means a DNA sequence capable of maintenance and replication within a host organism. The term "vector" includes, but is not limited to, plasmids such at pBluescript (Stratagene Inc., La Jolla, CA) or bacteriophages such as Lambda UniZAP (Stratagene).

ii) Probe Preparation

The DNA used to generate probes for subsequent arraying may be obtained from a large number of sources. For example DNA fragments may be obtained from a random selection of clones from a DNA library prepared from the organism of interest. In the case of animals such as man or rodents, these clones would preferably be obtained from one or more cDNA libraries. The fragments may also be selected from collections of clones which have been characterised to some extent, for example by partial sequence analysis of the insert DNA or by mapping of the insert DNA to particular chromosomal loci. Such clones may include, but are not limited to the I.M.A.G.E Consortium collection of clones isolated from human or rodent cDNA libraries and characterised by the generation of one or more ESTs for each clone [Lennon, G. et al., Genomics, 1996, 33, 151-152]. In the case of probes derived from bacterial genes, genomic DNA libraries may also be used. Each clone consists of a polynucleotide DNA fragment inserted at a known site within a suitable vector. The vector may for example be a plasmid vector such as pBluescript (Stratagene) or a bacteriophage vector such as Lambda UniZAP (Stratagene).

Individual bacterial clones from selected DNA libraries are cultured in the appropriate liquid medium using standard techniques. A small sample of each culture is used as a source of template DNA for subsequent amplification by PCR, using oligonucleotide primers

to a solid support at a specific physical location which defines one point within a 2-dimensional matrix constructed from a plurality of such elements.

As used herein, the term "EST" or Expressed Sequence Tag" refers to a partial DNA or cDNA sequence, typically of between 50 and 500 sequential nucleotides, obtained from a genomic or cDNA library prepared from a selected cell, cell type, tissue or tissue type, organ or organism which longer sequence corresponds to an mRNA of a gene found in that library [cf. Adams, M. D. et al., Science, 1991, 252, 1651-1656 and International Application No. PCT/US92/05222, published 7 January 1993]. An EST is generally DNA.

As used herein, the term "gene" refers to the genomic nucleotide sequence from which a cDNA sequence is derived.

As used herein, the term "immobilised" refers to the attachment of probe DNA to a solid support. The attachment may be of a covalent or non-covalent nature and will depend on the nature of the solid support being used.

As used herein, the term "insert" refers to any DNA sequence incorporated within a vector using methods of molecular biology available to anyone ordinarily skilled in the art.

As used herein, the term "oligonucleotide" refers to a molecule containing up to 50 nucleotides, but more typically 20 nucleotides of either DNA or RNA.

As used herein, the term "organism" includes without limitation, microbes, plants and animals.

As used herein, the term "probe" means a DNA species immobilised to a solid support within a DNA element.

As used herein, the term "solid support" refers to any known substrate which is useful for the immobilisation of probe DNA by any available method to enable detectable hybridisation of the immobilised oligonucleotides or polynucleotide DNA sequences to other polynucleotide sequences in a sample. Such useful solid supports include, but are not limited to, paper, nitrocellulose, myelin, glass, silica, nylon, plastics such as polyethylene, polypropylene or polystyrene, or other solid material. In addition, the term "solid support" can refer to gels constructed from such materials as agarose, polyacrylamide, polysaccharide or proteins, which may themselves be overlaid on a further solid surface such as glass or metal, to provide mechanical strength, electrical conductivity or other desired physical

precise quantitation of the absolute abundance of multiple cDNAs can be obtained within a single experiment.

Such single-strand arrays can be readily used to quantify the abundance of single-strand nucleic acid species such as mRNAs or their corresponding first-strand cDNAs in a variety of cell types or populations. The abundance information thus obtained can be used to draw up a quantitative transcript profile describing the expression of a large number of genes within any given cell type or cell population. This information can be used to determine for example which genes are differentially expressed in diseased versus normal tissue, or treated versus untreated tissue, and hence provide valuable information in diagnosing and monitoring disease processes, and in research to identify new treatments to restore the healthy state.

The invention will now be illustrated but not limited by reference to the following detailed description and Example:

i) Definitions

As used herein, the term "animal" is used in its broadest sense to include all members of the animal kingdom.

As used herein, the term "biological sample" encompasses any cell or tissue in any state from any organism which may be selected to provide a source of target nucleic acids.

As used herein, the terms "disease" or "diseased state" refer to any condition which deviates from the normal or standardised healthy state in an organism of the same species in terms of differential expression of the organism's genes. A disease state can be any illness or disorder of genetic or environmental origin which is characterised or may be described by the expression of genes which are either (i) normally silent in the healthy organism but activated in the diseased state as a cause of or in response to the disease, or (ii) normally expressed within some standard range in the healthy organism but over- or under-expressed in the diseased state as a cause of or in response to the disease.

As used herein, the terms "element" or "DNA element" refer to a number of immobilised DNA molecules, which may be either single-stranded or double-stranded, bound

The array conveniently comprises at least 10 DNA elements, such as at least 100 elements. Further convenient arrays comprise at least 1,000, 10,000 or 100,000 DNA elements.

The invention also provides a method whereby such selected single-strand arrays can be used to provide a quantitative estimate of the abundance of individual mRNAs or their corresponding first strand cDNAs within a complex mixture of such derived from a biological sample comprising a single cell type or a mixed population of cell types. The abundance is determined by measuring the amount of hybridisation between single-strand probe DNA at each element in the array and its complementary strand within the complex mixture of mRNA or cDNA. To detect such hybridisation, a label is incorporated into the mRNA or cDNA molecules in the complex mixture, for example a fluorescent nucleotide. In practising this aspect of the invention, mRNAs are isolated from cells and either directly labelled in vitro or, in a preferred embodiment, converted into first strand cDNAs, in which case the label is introduced on modified nucleotide which is incorporated into the single-strand cDNAs by reverse transcriptase. In this respect, the abundance of any given cDNA species within the population of single-strand cDNAs generated by reverse transcriptase is taken to represent the abundance of the corresponding mRNA within the biological sample.

The labelled cDNAs are hybridised to selected single-strand arrays which contain pairs of elements in which either the sense or antisense strand of each of the polynucleotide probes is immobilised at each element. The amount of immobilised DNA present at each element in the array is controlled such that it is considerably greater than the amount of the corresponding target mRNA or cDNA within the sample applied to the array. Under such conditions, the amount of labelled target nucleic acid (mRNA or cDNA) that remains bound to each element under the hybridisation conditions employed will represent the concentration of each mRNA or cDNA in the original sample. The bound target nucleic acid can be determined using an appropriate detection system capable of measuring the label carried on the target nucleic acid; e.g., a scanning fluorescence microscope [see e.g., Schena M. et al., Science, 1995, 270, 467-470]. The abundance of a particular cDNA (and hence its parental mRNA) may be quantified by comparing the intensity of the specific hybridisation signal, such as fluorescence intensity, at a given sense element to the non-specific hybridisation determined by the signal obtained at its corresponding antisense element. In this way, a

desired chemical modification(s) are selectively incorporated into the sense and/or antisense strands of the double-stranded DNA.

The primer(s) may be modified at any convenient position(s). Modification(s) are preferably made to the 5' nucleotide of the primer at either the 5' phosphate, the 5' deoxyribose group or the 5' base (adenine, guanine, thymidine or cytosine). In general, the modification involves the addition of a chemical functionality for binding to the solid surface, together with an optional spacer group of appropriate length to improve the accessibility of the probe DNA to the target nucleic acid. Both covalent and non-covalent binding may be used. In one embodiment, the chemical functionality may direct non-covalent binding to the solid surface, for example a biotin moiety which will interact with a streptavidin coating on the solid surface. In an alternative embodiment, the chemical functionality may covalently link the selected DNA strand to the solid surface.

Chemical modification of the DNA may be performed in one or more steps.

It will be appreciated that the sense and antisense strand of each DNA probe may be separated either prior to or after arraying onto the solid support. The separation may involve physical denaturation of the probes using for example heat or alkali, or the enzymic degradation of the unwanted strand for example using an appropriate exonuclease, or a combination of both methods.

In a further aspect of the invention we provide an array of single-stranded probe DNAs, each probe comprising at least 75 nucleotides and immobilised on a solid support. More conveniently each probe comprises at least 100 or 200 nucleotides, such as at least 500 or 1,000 nucleotides. A particular range is 300-10,000 nucleotides. A particular advantage of such an array is that the sequence of the probe DNAs may be unknown. Each probe DNA may be the sense or antisense strand for a given gene sequence. In further particular aspects of the invention every probe DNA in the array is antisense strand DNA or every probe DNA in the array is sense strand DNA.

the corresponding target DNA hybridising to a given DNA element, this is not a problem. However, there are applications where it would be advantageous to detect only one strand of any given target DNA in a complex mixture of target nucleic acids. For example, direct detection of a labelled mRNA target requires only an antisense strand DNA probe in each element. Alternatively, direct detection of a first-strand labelled cDNA target, synthesised from an mRNA template by reverse transcriptase, requires only a sense strand DNA probe. In such instances, the presence of the unwanted sense or antisense strand of the probe within each DNA element will reduce signal sensitivity by reducing the number of probe sites available for target hybridisation. It may also increase background signals by hybridising to non-specific target DNAs.

We have now devised methods for preparing single-strand arrays containing elements comprising either sense or antisense polynucleotide DNA probes. These are used to increase the sensitivity of the arrays when used as probes in hybridisation assays with either labelled RNA or labelled single-strand cDNA.

In a first aspect of the invention we provide a method for preparing an array of single-stranded DNA immobilised on a solid support, which method comprises (i) providing samples of double-stranded DNA chemically modified on the sense or antisense strand for attachment to the solid support, and (ii) linking the DNAs to the solid support and, before or after step (ii), removing the non-modified strand whereby an array of single-stranded DNA is provided on the solid support.

In a further aspect of the invention, a second chemical modification is provided on the strand that is not to be bound to the solid support. The purpose of this second chemical modification is to assist in either the separation of the two strands or the selective degradation of the unwanted strand.

The single-stranded DNA preferably comprises DNA molecules containing more than 75 nucleotides such as more than 100 nucleotides or more than 200 nucleotides. Preferred ranges of nucleotides include 100-10,000; 200-10,000 and 300-10,000.

The samples of double-stranded DNA chemically modified on the sense and/or antisense strand are conveniently provided by extension of chemically modified primer(s). Such primer(s) are preferably used as polymerase chain reaction (PCR) primers, whereby the

given element in the array is a measure of the concentration of the corresponding complementary cDNA in the original complex mixture [Schena M. et al., Science, 1995, 270, 467-470; Shalon, T. D. and Brown, P. O., International Patent Application No. WO 95/35505, published 28 December 1995; Pinkel, D. et al, International Patent Application No. WO 96/17958, published 13th June 1996].

Arrays of immobilised oligonucleotides have been described which use elements containing selected sense, antisense, missense or nonsense sequences at different positions in the array. Such arrays have been used for a number of applications; for example to determine the sequence of DNA [see e.g., Mirzabekov, A. D., Trends in Biotechnology, 1994, 12, 27-32 and references therein; Fodor, S. P. A. et al., International Patent Application No. WO 92/10588, published 25 June 1992; Chee, M. et al., International Patent Application No. WO 95/11995, published 4 May 1995]. In another application, arrays of allele-specific oligonucleotides have been used to detect genetic polymorphisms and determine genotypes [see e.g., Southern, E., European Patent No. 0373203B1, published 31 August 1994; Guo, Z. et al., Nucleic Acids Research, 1994, 22, 5456-546]. However, such arrays have limitations when used to probe complex mixtures of labelled polynucleotide DNA targets such as cDNAs, since any given oligonucleotide may hybridise at a particular stringency to sequences in more than one target DNA. In the case where the sequence of the target DNAs are known, it is possible to design sets of oligonucleotides which will provide a unique hybridisation signal for each gene. Such sets can be combined into one or more elements of an array to provide a hybridisation signal characteristic for any known target [Fodor, S. P. A. et al., International Patent Application No. WO 92/10588, published 25 June 1992; Chee, M. et al., International Patent Application No. WO 95/11995, published 4 May 1995]. However, such an approach cannot be accurately applied where the sequence of the target DNAs is unknown or incomplete. Where the oligonucleotide arrays are constructed by in situ synthetic methods, the addition of an additional target gene requires the whole array to be resynthesised, with considerable cost implications.

In comparison, arrays of longer non-oligonucleotide probe DNAs provide a much higher specificity for hybridisation to target DNAs. However, all such arrays to date have incorporated a mixture of both sense and antisense strands of a particular DNA fragment within each DNA element in the array. In cases where it is desired to detect both strands of

et al., International Patent Application No. WO 95/11995, published 4 May 1995]. An alternative approach has been described by Southern [Southern, E., European Patent No. 0373203B1, published 31 August 1994]. The maximum length of oligonucleotides assembled in these arrays is restricted by the chemistry employed to assemble the oligonucleotides, which in practice are usually no more than 20-25 nucleotides in length.

In another approach, arrays of longer DNA species may be constructed by using robotic micropipetting devices to transfer small, typically nanolitre or smaller quantities of DNA from containers such as 96-well plates to ordered pre-determined positions on a non-porous surface such as a glass microscope slide Each DNA sample is bound at a known position on the microscope slide to constitute one DNA element of the array. Using such apparatus a large number of replica slides can be constructed supporting arrays of thousands of individual DNA elements [Schena M. et al., Science, 1995, 270, 467-470; Shalon, T. D. and Brown, P. O., International Patent Application No. WO 95/35505, published 28 December 1995].

In this latter approach, the DNA samples being transferred to the solid support are typically double-stranded polynucleotide DNA fragments of length greater than 50bp. These DNAs may be obtained from a number of sources, such as cDNA or genomic DNA libraries, and may be of either known or unknown sequence composition.

DNA may be coupled to the solid support by a number of techniques. For example, the DNA may be bound to glass through non-covalent electrostatic interactions with a coating film of a polycationic polymer such as poly-L-lysine [see e.g. Shalon, T. D. and Brown, P. O., International Patent Application No. WO 95/35505, published 28 December 1995]. Alternatively, DNA can be bound covalently to the solid support. There are a number of methods available for covalent linkage of DNA to solid supports, depending on the nature of the support.

Arrays of polynucleotide DNA probes immobilised on solid supports can be used to study the composition of complex mixtures of DNA using hybridisation techniques. In a typical application, a complex mixture of labelled cDNA is hybridised to the DNA array under conditions of appropriate stringency, and unbound material is washed away. The array is then scanned using a detection method capable of sensing the remaining bound labelled cDNA, such as a scanning fluorescent microscope. The intensity of the detected signal at any

2318791

METHODS

This invention relates to methods for preparing arrays of nucleic acids for use in biological screening procedures such as hybridisation assays, with applications in genetic research and diagnostic applications.

Increasing use is being made of arrays of immobilised nucleic acids, particularly arrays of DNA, for genetic research and diagnostic purposes. These arrays consist of a plurality of DNAs organised as a two-dimensional matrix immobilised on an appropriate solid support. Each point in the matrix comprises a DNA element. Each of the DNA elements can be used as a probe to detect complementary sequences in complex mixtures of nucleic acid. This allows parallel determination of the identity and abundance of many DNA species in a single experiment.

Such arrays can be formed on porous membranes such as nitrocellulose using a variety of methods. In the dot-blot or slot-blot technique, a plurality of DNA samples is transferred to membranes by placing the samples into a manifold consisting of an array of pre-formed wells applied to the top of the membrane, and drawing the DNA through the membrane using a vacuum. In another variant of this method, DNA is applied directly to the membrane using an array of pins to transfer DNA onto the membrane surface from DNA samples contained, for example, in the wells of a microtitre plate [Lehrach, H. et al., "Hybridisation Fingerprinting in Genome Mapping and Sequencing" in "Genome Analysis", Vol. 1, Davies, K. E. and Tilghman S. M. (Eds.), Cold Spring Harbor Laboratory Press, New York, 1990, pp38-82; Nizetic, D. et al., Proceedings of the National Academy of Sciences (USA), 1991, 88, 3233-3237].

Alternatively, DNA arrays can be formed on non-porous surfaces such as glass, by either in situ synthesis or direct application. For example, arrays of oligodeoxynucleotides can be assembled by starting with a chemically sensitised glass surface which is protected by a mask, and reacting selected exposed areas with suitably modified nucleotides. By appropriate choice of masks and nucleotide reagents, arrays of synthetic oligodeoxynucleotides of defined sequence can be elaborated at the glass surface [see e.g., Jacobs, J. W. and Fodor, S. P. A., Trends in Biotechnology, 1994, 12, 19-26; Fodor, S. P. A. et al., International Patent Application No. WO 92/10588, published 25 June 1992; Chee, M.

(12) UK Patent Application (19) GB (11) 2 318 791 (13) A

(43) Date of A Publication 06.05.1998

(21) Application No 9722594.0

(22) Date of Filing 28.10.1997

(30) Priority Data
(31) 9622665.9 (32) 31.10.1996 (33) GB

(71) Applicant(s)
Zeneca Limited

(Incorporated in the United Kingdom)

15 Stanhope Gate, LONDON, W1Y 6LN, United Kingdom

(72) Inventor(s)
Andrew David Charles

(74) Agent and/or Address for Service
Neil Godfrey Alasdair Phillips
Zeneca Pharmaceuticals, Intellectual Property Group, Intellectual Property Department, Mereside, Alderley Park, MACCLESFIELD, SK10 4TG, United Kingdom

(51) INT CL^6
C12Q 1/68 , C07H 21/04

(52) UK CL (Edition P)
C3H HB5
G1B BAG B121 B309 B500
U1S S1334 S1337

(56) Documents Cited
WO 95/35505 A1
Science 1995,270,467-470 Nucleic Acids Research 1994,22(24),5456-5465 Analytical Biochemistry 1993,209,63-69

(58) Field of Search
UK CL (Edition O) C3H HB4A HB4B HB5 , G1B BAG BAH
INT CL^6 C12Q 1/68
ONLINE: WPI,CLAIMS,DIALOG/BIOTECH

(54) Array of single-stranded DNA immobilised on a solid support

(57) An array of single-stranded DNA probes, each of which comprises at least 75 nucleic acid units, is chemically immobilised on a solid support. The probes may be used to provided a quantitative estimate of the abundance of individual mRNA (or the first strand cDNA corresponding thereto) within a complex mixture thereof, obtained from a biological sample comprising a single cell type or a mixed population of cell types. The probe DNA may be of unknown sequence and may comprise antisense, or sense, strand DNA.

An array of single-stranded DNA, immobilised on a solid support, is prepared by:

(i) provision of samples of double-stranded DNA chemically modified on the sense, or antisense, strand for attachment to the solid support;

(ii) linking the DNA to said support:

wherein, prior to, or after, (ii), the non-modified strand is removed.

The strand, not bound to the support, may be chemically modified, either to assist strand separation or the selective degradation thereof. The double-stranded DNA may be the amplification products of chemically modified primers obtained by the polymerase chain reaction.

GB 2 318 791